Applications of the Fluctuation Theorem

Emil Joshua Mittag

B.Sc., M.Sc., University of Texas at El Paso, USA

March 2002

THE AUSTRALIAN NATIONAL UNIVERSITY

A thesis submitted for the degree of Doctor of Philosophy of

The Australian National University.

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person, except where due acknowledgement is made in the text of the thesis.

Acknowledgements

I would like to thank Professor Denis J. Evans for allowing me to join his research group, for his guidance and for his supervision of the work outlined in this thesis during my Ph.D. studies. I am also grateful to him for giving me the chance to participate in some excellent bushwalks and to go snowshoeing in the Australian Alps. I would also like to thank Dr. Debra J. Bernhardt for her supervision and for spending so much of her time writing e-mail to answer my sometimes incessant questions. Her help in debugging some of the code used in simulations performed during the research for this thesis is also gratefully acknowledged. I would like to thank Dr. Janka Petravic for her help in explaining some of the concepts of non-equilibrium statistical mechanics and for always being interested in what I was doing. Prof. Dr. Siegfried Hess is gratefully acknowledged for his helpful discussions and suggestions during his visit to the RSC. I thank Professor Leo Radom for helping me come to Australia for my Ph.D. studies.

I must also thank the computer support unit of the RSC (Pam Cohen, Dr. Piotr Wielopolski, Dr. Graeme Lindsell, Cliff Dixon and Dr. Christopher Delfs) for keeping the LAN running so I could do research and write my thesis. I would also like to acknowledge the help of Dr. David Singleton and Dr. Ben Evans of the Australian National University Supercomputer Facility for their help in answering my questions when I first parallelised some code using MPI.

The Department of Computer Science at ANU is gratefully acknowledged for employing me as a course tutor over the last two years. During that time I not only taught many interested students, but also learned a great deal. I would like to thank Dr. Michael Hartmann for his advice during the brief time he was at the RSC and for his continuing friendship. I thank Richard Walker for many enlightening philosophical discussions over coffee. I also want to thank Dr. Wolfram Wielandt for his friendship while he was at the RSC and for his time spent practising Spanish with me; without his help, I surely

would have forgotten how to speak it.

I wish to thank Professor William C. Herndon, Professor Keith Pannell and Professor M. Lawrence Ellzey, all of the University of Texas at El Paso, for their advice and guidance during my first two degrees. Their help was invaluable and I could not have come this far without it.

Finally, I thank my wife, Joanna, for being patient and supportive during my undergraduate and postgraduate studies.

Lernt es ruhig mal so, ich verstehe auch nicht alles was ich weiß.

Reinhold Löffler

This is by no means the only occasion where philosophical thought entangles itself in contradictions...It finds insurmountable difficulties in the relationship between cause and effect, between body and soul, in the possibility of consciousness, in short, in all and everything...in the end it finds it wholly inexplicable and self-contradictory that anything at all exists, that anything has arisen and is capable of change.

Ludwig Boltzmann

What is mind-boggling is the power of computers, the power we use to play sorcerer's apprentice by writing a deceptively small software text and then executing it to create objects and perform computations on them in numbers so large—number of objects, number of computations—as to appear almost infinite when measured on the scale of human understanding.

Betrand Meyer

Abstract

The Fluctuation Theorem (FT) is an analytical expression for the probability that the Second Law of Thermodynamics may be violated in a finite system observed for a finite time. This thesis describes the theoretical basis of the FT and reports on several new applications of this theorem. Molecular dynamics simulations have been designed which test the Fluctuation Theorem in several ensembles and the results of these tests are reported. Various dynamics are considered, including constant temperature, constant temperature–constant pressure and adiabatic. The results of a recent laboratory experiment that verifies the FT and molecular dynamics simulations that model this experiment are presented. Finally, the existence of a time-dependent FT is demonstrated and numerical results are presented to confirm theoretical predictions.

Contents

List of Figures

Chapter 1

Introduction

Computers and Statistical Mechanics

Since the 1960s, computers have been used to study the behaviour of various physical phenomena. These phenomena include chemical reactions, the structure of matter, the behaviour of solids and liquids, nuclear reactions, and biological systems, to name but a few. Initially, the calculations that were performed were primitive by modern standards, as the computer hardware at that time was extremely expensive and much less powerful than contemporary computer systems. As Moore's Law* continues to hold, computers have become much more powerful than those used in initial simulations of physical phenomena. Modern desktop machines are able to outperform the fastest supercomputer of only a decade ago, and contemporary pocket calculators are more powerful than the fastest machines in the 1960s. This increase in computational power increased the ability to study many physical phenomena via computer simulations.

Statistical mechanics is the study of macroscopic systems through the study of individual particles, whether they be atoms or molecules. Statistical me-

*Moore's Law has its origin in an observation made by Gordon Moore in 1965. At that time Moore noticed that the number of transistors per square inch on an integrated circuit had doubled every year since the integrated circuit had been invented. He predicted that this trend would continue in the foreseeable future. Currently, it is still generally regarded that Moore's Law continues to hold. The ratio, however, is approximately double the number of transistors per 1.5 years.

chanics allows one to calculate macroscopic properties of a given system from a knowledge of the molecular behaviour. These systems may be at equilibrium, or they may be subjected to some external force that creates a non-equilibrium state. Statistical mechanics predicts the properties or states of a system without exact knowledge of an initial starting state. This contrasts with ordinary mechanics, where one has a precise knowledge of an initial state and may proceed from there to another exactly known state. In statistical-mechanical studies conducted on liquids, inter-particle interactions are very important, while in the ideal gas (density $\rightarrow 0$) they are less so. For liquid states, much work has therefore been done on formulating and improving model interaction potentials. The application of statistical mechanics via the solution of the equations of motion is referred to as molecular dynamics simulation, since the behaviour of each particle (atom, molecule) is considered.

The study of what comprises contemporary statistical mechanics stems from the work of Maxwell, Boltzmann and Gibbs in statistical thermodynamics (see, for example, [1–3]). This early work on statistical mechanics dealt with systems both at and away from equilibrium. Much of the contemporary work in non-equilibrium statistical mechanics is based on research which was done well into the 20^{th} century (see, for example, [4–8]), indicating that the field of non-equilibrium statistical mechanics is indeed young.

Since many of the concepts in non-equilibrium statistical mechanics are relatively new, there is still much that is not understood. Hoover [9], and separately Evans and Morriss [10], published works on non-equilibrium molecular dynamics (NEMD) which offered new insight into non-equilibrium phenomena and new mathematical and computational techniques for studying non-equilibrium systems.

The methodology used to study statistical mechanics has changed as the technology available to researchers has improved. It is not just the speed of computers that have fundamentally changed the way in which research is carried out in this field, but also the mere *possibility* of performing numerical experiments. Early research in this field was comprised of either developing

theories that would explain phenomena observed in experimental laboratories or using mathematical reasoning to predict some phenomenon that was later verified by experiment. Often if the results of experiments did not agree with the theory, it was not clear if the model potential or underlying theory was wrong.

The approach to doing research in statistical mechanics has changed dramatically with the advent of computers. Contemporary research in this field often entails the development of new theories and the subsequent testing of these theories via computer simulations. It is less common that actual laboratory experiments are carried out to verify theories for which numerical simulations can be devised. The ambiguity of experimental results not agreeing with the underlying theory has been removed; computer simulations can be carried out on model systems identical to those used in the underlying theory.

This thesis documents recent investigations into problems of non-equilibrium statistical mechanics and the use of computer simulations as tools in helping to solve these problems, with particular emphasis on investigations associated with the Fluctuation Theorem. The Fluctuation Theorem is a derivation of the Second Law, which quantifies in small systems observed for short times, fluctuations in the sign of the entropy production. Before discussing the nuances of the Fluctuation Theorem in detail in Chapter 3, the mathematical theory useful in statistical mechanics as applied to molecular dynamics simulations will be discussed in Chapter 2.

Chapters 4, 5 and 6 discuss recent studies involving the Fluctuation Theorem and show the numerical results of computer simulations in addition to theoretical derivations of alternate forms of the Fluctuation Theorem. Chapter 7 examines a recent laboratory experiment that verifies the Fluctuation Theorem and molecular dynamics simulations of this experiment. Chapter 8 discusses the validity of the Fluctuation Theorem for systems with time-dependent external fields.

Chapter 2

Molecular Dynamics

2.1 Introduction

Molecular dynamics (MD) involves the use of computers to simulate the movement of atoms and molecules by solving Newton's equations of motion. The aim of such simulations is to compute the macroscopic properties (e.g. temperature, pressure) of a system. Simulations have gained in popularity as the scope of systems that may be adequately treated in this manner has increased due to faster computer hardware and better algorithms. Molecular dynamics is commonly used in liquid-state simulations to obtain macroscopic properties of liquids in the presence of an external field. The techniques associated with these non-equilibrium simulations are referred to as non-equilibrium molecular dynamics (NEMD). An advantage of computing macroscopic properties with molecular dynamics is that MD simulations may be done on systems for which laboratory experiments are difficult, expensive or impossible.

The first MD simulation was performed by Alder and Wainwright in 1957 [11]. Much work has been done in improving computer simulation methods since that time [11,12]. Research in the areas of numerical analysis, statistical mechanics, algorithms and complexity and fluid dynamics has been applied to liquid-state MD simulation theory.

Molecular dynamics simulations normally arrive at approximate results. Causes of error include truncation error, discretisation error, errors due to

finite N (the number of particles in the simulation) and errors due to the approximate nature and form of the interparticle potential. Truncation error is a problem intrinsic to calculations by computer. Computers have finite numbers of decimal places with which they represent real numbers and calculate. If one uses too few decimal places in ones calculations, the results will not be accurate. If one uses more decimal places, the computational cost of the calculations may increase dramatically.

Discretisation error is caused by the fact that computers can only solve difference equations. An ordinary differential equation (ODE) must therefore be approximated by a difference equation. The computer solution of an ODE has the form

$$\text{ODE} = \text{difference equation} + \mathcal{O}(\triangle t^n),$$

where the magnitude of the discretisation error is characterised by n, which typically has a value of 3 or 4 in common algorithms.

Errors due to finite N are reduced but not eliminated by the use of periodic boundary conditions. Sometimes only a small number of particles is necessary to investigate the physical phenomenon in which one is interested, while other simulations require thousands or even tens of thousands of particles to yield reliable or reasonable results. Often it is possible to examine the theory associated with a planned MD simulation to determine an approximate number of particles necessary to achieve good results. In order to determine the proper number of particles for a given simulation, one may perform test simulations to observe how the results converge (if they do so) as the number of particles increases.

The approximate nature of the interparticle potential can result in errors, especially when compared to experimental results. The magnitude of this error can be substantially increased by the choice of the thermostatting mechanism and the equations of motion used.

2.2 Mechanics

It is well known that Newton's equations,

$$\mathbf{F} = m\mathbf{a} = m\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}, \tag{2.1}$$

may be used to describe the motion of classical bodies.

For the sake of simplicity and also for the solution of some specific types of problems, it is desirable to find a form of the equations of motion that is in terms of the coordinates and velocities and is valid in any coordinate system. Generalised coordinates may be defined whereby both of these conditions may be satisfied. Generalised coordinates allow for the treatment of systems where coordinates can be eliminated by applying constraint equations without explicitly considering the forces of constraint. Hamiltonian mechanics (§2.2.1) provides a framework for a generalised coordinate system, which is much more flexible than Newtonian mechanics. The canonical coordinates given in Hamilton's equations are also the coordinates that should be used in Gibbs prescription for equilibrium statistical mechanics.

2.2.1 Hamiltonian Mechanics

Hamiltonian mechanics is based on work originally done by Lagrange and later reworked by Hamilton. Lagrange defined what is now known as Langrange's equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{q}}_i}\right) = \frac{\partial\mathcal{L}}{\partial\mathbf{q}_i} \tag{2.2}$$

and the corresponding *Lagrangian*, $\mathcal{L}(\mathbf{q}_j, \dot{\mathbf{q}}_j, t)$; $j = (1 \ldots f)$ where f is the number of degrees of freedom in the system. Note that here the $\mathbf{q}_j$'s and the $\dot{\mathbf{q}}_j$'s are the generalised coordinates and velocities of the system. The Lagrangian is defined by $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = T - V$, where $T(\mathbf{q}, \dot{\mathbf{q}})$ is the kinetic energy and $V = V(\mathbf{q})$ is the potential energy.

For a system that uses rectilinear coordinates, such as the Cartesian coordinate system, $\mathcal{L} = T - V$, where the kinetic energy, T is defined by

$T = \sum_i \frac{1}{2} m_i \dot{\mathbf{x}}_i^2$ and the potential energy is $V = V(\mathbf{x}_i)$. It then follows that $\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{x}}_j} = m\dot{\mathbf{x}}_j \equiv \mathbf{p}_j$, which is the component of the momentum in the j direction. Expressing this in terms of *generalised* momenta, we have $\mathbf{p}_j = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}_j}$, which relates the generalised momenta to the generalised coordinates $\mathbf{q}_j$.

If the Lagrangian of a system is not explicitly time-dependent, then the energy, E, of the system will be constant. The value of the energy of a given system which has Newtonian or Lagrangian equations of motion is equal to the sums of the potential and kinetic energies associated with that system. Using the notation of Lagrangian mechanics, we have

$$\mathsf{E} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}_i} \dot{\mathbf{q}}_i - \mathcal{L}. \tag{2.3}$$

If we define $\mathcal{H}(\mathbf{q}, \mathbf{p}) \equiv \mathsf{E}$ for the system, and use the relation $\mathbf{p}_i = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}_i}$, we have

$$\mathcal{H}(\mathbf{q}, \mathbf{p}) = \sum_i \mathbf{p}_i \dot{\mathbf{q}}_i - \mathcal{L}. \tag{2.4}$$

This allows the definition of what are known as Hamilton's equations

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\partial \mathcal{H}(\mathbf{q}, \mathbf{p})}{\partial \mathbf{p}_i} \\ \dot{\mathbf{p}}_i &= -\frac{\partial \mathcal{H}(\mathbf{q}, \mathbf{p})}{\partial \mathbf{q}_i} \end{aligned} \tag{2.5}$$

where $\mathcal{H}(\mathbf{q}, \mathbf{p})$ is the system *Hamiltonian*; the Hamiltonian is defined in terms of the generalised coordinates and momenta and the time of the system.

The equilibrium Hamiltonian for point particles is defined as

$$\mathcal{H}_0(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{\mathrm{N}} \frac{\mathbf{p}_i^2}{2m} + \Phi(\mathbf{q}). \tag{2.6}$$

Note that the inter-particle potential energy, $\Phi(\mathbf{q})$ is a function of the generalised coordinates of all particles. Using Hamilton's equations, we can derive

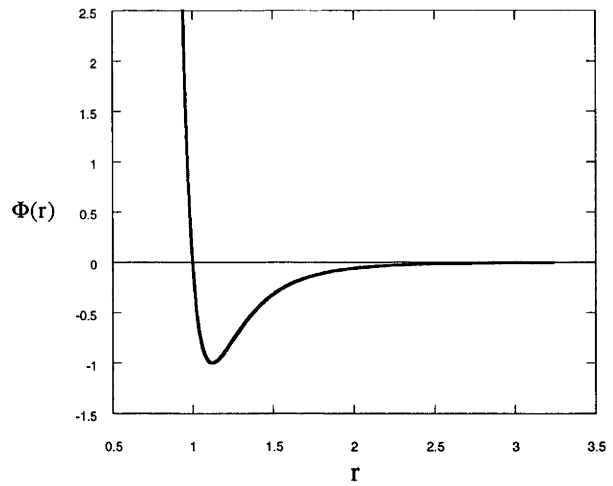


Figure 2.1: The Lennard-Jones 12-6 pairwise interaction potential, shown with $\epsilon = 1$ and $\sigma = 1$.

equations of motion for a system

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i. \end{aligned} \tag{2.7}$$

$\mathbf{F}_i$ is the force on particle i, which is defined by $\mathbf{F}_i \equiv -\frac{\partial \Phi(\mathbf{q})}{\partial \mathbf{q}_i}$. These equations are those most commonly used in molecular dynamics to simulate particle motion. Often these equations of motion are modified to produce alternate dynamics where quantities such as the temperature (§2.9) or pressure (§6.2.1) are constants of the motion.

2.3 Interaction Potentials

As alluded to above, to calculate the forces on particles in a MD simulation, one must define an interaction potential. Much work has been done on developing such potentials [11]. One of the best-known potentials is the Lennard-Jones (LJ) potential, which may be divided into attractive and repulsive parts. The mathematical form of the 12-6 Lennard-Jones potential is

$$\Phi(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{2.8}$$

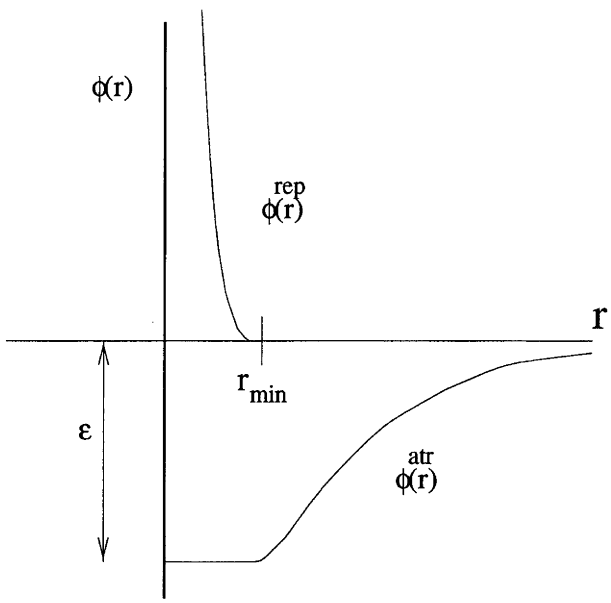


Figure 2.2: The Lennard-Jones interaction potential, shown divided into its attractive and repulsive parts.

where ϵ is the well depth at the minimum of $\Phi(r)$ and σ is a parameter characterising the separation of particles and the value of r at which the potential changes sign (i.e. $\Phi(r) = 0$). Figure 2.1 shows the LJ potential with $\epsilon = \sigma = 1$, a common condition for MD simulations in *reduced units*, which are discussed in greater detail in (§2.7.4).

As r increases, the attractive part of the LJ potential gradually converges to zero. For this reason, it is common to set a cutoff value, r_{cut}, beyond which any inter-particle interactions are ignored. One must use care in setting this value, as if it is too short, important interactions will be ignored, while if it is too long, the efficiency of the MD algorithm will suffer. For the Lennard-Jones potential, the cutoff value is commonly set to 2.5σ. [11] With a cutoff, the LJ potential may be rewritten

$$\Phi(r) = \begin{cases} 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] & r < r_{cut}, \\ 0 & r > r_{cut}. \end{cases} \tag{2.9}$$

It is possible to divide the Lennard-Jones potential into its repulsive and attractive parts, splitting it at the minimum value of the potential, $r = r_{min}$. The Lennard-Jones repulsion between two particles is modelled by

$$\phi^{rep}(r) = \begin{cases} \Phi(r) + \epsilon & r < r_{min} \\ 0 & r_{min} \leq r, \end{cases} \tag{2.10}$$

where the repulsion has been shifted up by an amount ϵ so the potential is zero at the cutoff value for the repulsion. The LJ attraction is modelled by

$$\phi^{atr}(r) = \begin{cases} -\epsilon & r < r_{min} \\ \Phi(r) & r_{min} \leq r. \end{cases} \tag{2.11}$$

In the Lennard-Jones potential r_{min} is equal to $2^{\frac{1}{6}}\sigma$. A diagram illustrating the attractive and repulsive parts of this potential is shown in Figure 2.2.

2.3.1 WCA Potential

The Weeks-Chandler-Andersen potential [13] (WCA) is another commonly used pairwise additive interaction potential in MD simulations. This potential is a modified Lennard-Jones potential, which has the form given in Equation (2.10). Note that this form of the WCA potential also includes a shift of ϵ for the repulsion. All inter-particle interactions that take place at distances greater than r_{min} are ignored. Therefore, for the WCA potential, $r_{cut} = r_{min} = 2^{\frac{1}{6}}\sigma$.

The WCA potential is useful for systems where attractive interactions are not important. While the LJ potential is obviously more realistic than that of WCA, WCA is sometimes more useful for testing theorems of statistical mechanics. When testing such theorems, such as the Fluctuation Theorem, emulation of a real system is not necessary. For these types of problems, it is more important for the simulation to run faster, rather than for the results to be comparable to a real system. The WCA potential is much faster than LJ due to the lack of interactions beyond r_{min} [14].

2.4 Preliminary Considerations

Before discussing the nuances of molecular dynamics implementations of the principles of statistical mechanics, it is necessary to mention some concepts fundamental to both statistical mechanics and molecular dynamics. These include ensemble averaging, the use of external fields and thermostats.

In statistical mechanics, one is interested in properties averaged over a large number of microstates. Each of these microstates satisfies the macro conditions of the system. A large (in principle, infinite) number of such microstates is called an *ensemble.*

2.4.1 External Fields

In order to force an equilibrium system away from its thermodynamic equilibrium state, it is often necessary to apply an external field. This external field does work on the system, which has various effects. Among these effects is an increase in the system's temperature.

In the presence of an external field, the time derivative of coordinates and/or momenta will have a term coupling them to this field for each particle. The equations of motion (Equation (2.7)) then become

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} + \mathbf{C}_i\mathbf{F}_e \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i + \mathbf{D}_i\mathbf{F}_e \end{aligned} \tag{2.12}$$

where $\mathbf{C}$ and $\mathbf{D}$ are terms that couple the external field to the coordinates and momenta, respectively and $\mathbf{F}_i$ is the force on particle i.

2.4.2 Thermostats

In molecular dynamics simulations that incorporate an external field that does work on the system (§2.4.1), it is often desirable to maintain the temperature at a constant value. This is also required in simulations that generate certain equilibrium ensembles (e.g. the isokinetic and canonical ensembles). In order

to accomplish this, a new term is added to the equations of motion which maintains the system at a constant kinetic energy. In effect, this extra term removes heat from the system and functions as an infinite heat sink. The equations of motion (Equation (2.7)) with a term for the heat sink become

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \alpha \mathbf{p}_i. \end{aligned} \tag{2.13}$$

Further discussion on this topic and actual mathematical forms of two well-known thermostatting mechanisms are discussed in (§2.9).

2.5 Phase Space

For purposes of the following discussion, it is assumed that we are dealing with an N particle system that models a three-dimensional fluid. The discussion is, however, applicable to systems of dimensions other than three, such as two-dimensional models. One may define a 6N-dimensional space, the dimensions of which are comprised of 3N coordinates and 3N momenta. This space is known as the *phase space* of a system. Any given point in phase space is uniquely characterised by its coordinates and momenta. To simplify the discussion of phase space, we define

$$\mathbf{\Gamma} \equiv (\mathbf{q}, \mathbf{p}) \equiv (\mathbf{q}_1, \mathbf{q}_2, \ldots \mathbf{q}_N, \mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_N) \tag{2.14}$$

for a given 6N-dimensional point in phase space ($\mathbf{\Gamma}$-space). As a system evolves in time, a trajectory is traversed wherein the coordinates and momenta change (see Figure 2.3). In order to follow the evolution of a $\mathbf{\Gamma}$-space trajectory from one point in $\mathbf{\Gamma}$-space to the next, it is necessary to solve 6N first-order differential equations.*

*or 3N second-order differential equations, as in Newton's equations.

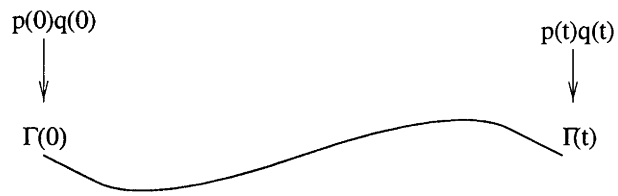


Figure 2.3: As a trajectory is propagated through phase space, the coordinates and momenta change with time. The trajectory is computed from the dN (where d is the dimension of the phase space) first-order differential equations of motion for the system.

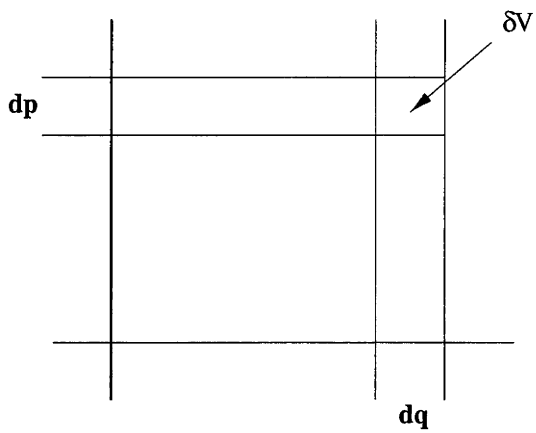


Figure 2.4: Illustration of a phase volume element, δV.

2.5.1 Phase-space Distribution and the Liouville Equation

In order to characterise the distribution of ensemble members in Γ-space, it is necessary to consider individual volume elements, as depicted in Figure 2.4. These infinitesimal volume elements, shown in the figure as $\delta\mathbf{V}$, are defined as

$$\delta\mathbf{V} = \lim_{\delta\mathbf{p},\delta\mathbf{q}\to 0} \delta\mathbf{q}_1\delta\mathbf{q}_2\cdots\delta\mathbf{q}_N\delta\mathbf{p}_1\delta\mathbf{p}_2\cdots\delta\mathbf{p}_N. \tag{2.15}$$

The phase-space distribution function is defined in terms of the phase-space volume and the probability of finding an ensemble member at $\mathbf{\Gamma}$ within this volume,

$$f(\mathbf{\Gamma},t) \equiv \frac{\Pr(\delta\mathbf{V})}{\delta\mathbf{V}}. \tag{2.16}$$

Since the number of ensemble members is constant, this phase-space distribution function can be normalised, i.e. $\int f(\mathbf{\Gamma},t)\mathrm{d}\mathbf{\Gamma} = 1$. The phase-space distribution function is therefore often referred to as a *probability* distribution

function.

With any given closed volume δV in $\mathbf{\Gamma}$-space, the number of ensemble members inside this volume is equal to the total flux of ensemble members into the volume,

$$f(q_1,\ldots)\dot{q}_1(q_1,\ldots)\delta q_2\ldots\delta q_{3\mathrm{N}}\delta\mathbf{p} \tag{2.17}$$

minus the total flux out of the volume,

$$f(q_1+\delta q_1,\ldots)\dot{q}_1(q_1+\delta q_1,\ldots)\delta q_2\ldots\delta q_{3\mathrm{N}}\delta\mathbf{p}. \tag{2.18}$$

These fluxes can be through any given face of the volume element normal to the q_1 and q_2 axes. Therefore, the probability of finding a trajectory segment in a given volume element will, in general, change with time. Summing over all phase-space directions gives [10]

$$\frac{\mathrm{dPr}(\delta\mathrm{V})}{\mathrm{d}t} = -\sum_{i=1}^{\mathrm{N}}\Big[f\Big(\frac{\partial}{\partial\mathbf{\Gamma}}\cdot\dot{\mathbf{\Gamma}}\Big)+\dot{\mathbf{\Gamma}}\cdot\frac{\partial f}{\partial\mathbf{\Gamma}}\Big]\delta\mathbf{q}\delta\mathbf{p} \tag{2.19}$$

since the change in Pr(δV) is dependent on fluxes in each of the $2dN$ directions.

Combining Equations (2.16) and (2.19) gives the relation for the rate of change in the probability density of the system

$$\frac{\partial(\frac{\mathrm{Pr}(\delta\mathrm{V})}{\delta\mathbf{q}\delta\mathbf{p}})}{\partial t} = \frac{\partial f}{\partial t} \tag{2.20}$$

which may be rewritten [10]

$$\frac{\mathrm{d}f}{\mathrm{d}t} = -f\frac{\partial}{\partial\mathbf{\Gamma}}\cdot\dot{\mathbf{\Gamma}} \equiv -f\Lambda(\mathbf{\Gamma}). \tag{2.21}$$

This is the Liouville equation and $\Lambda(\mathbf{\Gamma})$ is the phase-space compression factor.

As a trajectory is propagated through phase space, it is possible for the distribution about $\mathbf{\Gamma}$ to change. If we therefore consider a set of phase points in the volume δV at time zero, a volume $\delta\mathrm{V}(t)$ that is occupied by these points at a later time t will also change. As indicated above, the change in a given phase volume with time is related to $\frac{\mathrm{d}f}{\mathrm{d}t}$ and is characterised by $\Lambda(\mathbf{\Gamma})$.

For a Hamiltonian system in the presence or absence of an external field, there is no phase-space compression. Therefore, if a system has equations of motion that are Hamiltonian and one follows the streaming time evolution of phase space volume elements, these volume elements will not change in volume, although they do change in shape. Although equations of motion that are Hamiltonian are sufficient to guarantee $\Lambda = 0$, this is not a necessary requirement. As will be discussed later, it is possible to have equations of motion that are not derivable from a Hamiltonian, yet still have no phase-space compression.

In order to simplify the characterisation of the change in $\mathbf{\Gamma}$-space distribution about a point $\mathbf{\Gamma} \equiv (\mathbf{q}, \mathbf{p})$, it is possible to introduce a more compact notation for the Liouville equation. Equation (2.20) may be rewritten as

$$\frac{\partial f(\mathbf{\Gamma}, t)}{\partial t} = -\Big(\Big(\frac{\partial}{\partial \mathbf{\Gamma}} \cdot \dot{\mathbf{\Gamma}}\Big) + \dot{\mathbf{\Gamma}} \cdot \frac{\partial}{\partial \mathbf{\Gamma}}\Big) f \equiv -iL(\mathbf{\Gamma}) f(\mathbf{\Gamma}, t). \tag{2.22}$$

The distribution function Liouville operator, $-iLf$, allows one to state simply the relation between the $\mathbf{\Gamma}$-space distribution at two different times [10]

$$f(\mathbf{\Gamma}, t) = \mathrm{e}^{-iLt} f(\mathbf{\Gamma}, 0). \tag{2.23}$$

2.6 Equilibrium Ensembles

In molecular dynamics simulations and statistical mechanics in general, one is interested in studying the bulk properties of fluids. For such systems, typically the volume, V, is constant in addition to the number of particles, N. Often, the internal energy, E, or the temperature, T, is also held constant. Two ensembles, NVE and NVT, are discussed below.

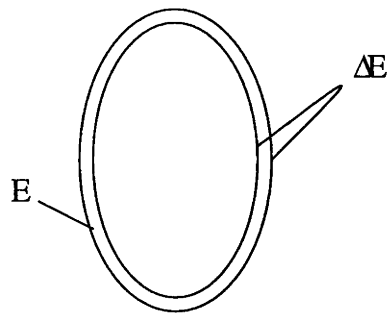


Figure 2.5: A depiction of constant microcanonical hypersurfaces.

2.6.1 Microcanonical Ensemble

In the microcanonical ensemble, groups of phase points with the same energy are considered. The microcanonical ensemble is often referred to as an NVE ensemble, as the number of particles, the volume and the energy of each system in the ensemble are held constant. When performing MD simulations in the microcanonical ensemble, we are interested in the thin-shell microcanonical ensemble, where the energy of the system is constrained to a value within a certain small tolerance. The probability distribution function for the thin-shell microcanonical ensemble is

$$f(\Gamma, \mathsf{E}, \Delta\mathsf{E}) = \frac{1}{\int_{\mathsf{E}<\mathcal{H}_0(\Gamma)<\mathsf{E}+\Delta\mathsf{E}} \mathrm{d}\Gamma} \tag{2.24}$$

where $\mathcal{H}_0(\Gamma)$ is the equilibrium Hamiltonian and E is a given energy [11]. In this mathematical description all solutions to the equations of motion that would result in a total energy of the system not within a given range are discarded. This constant energy thin shell is illustrated in Figure 2.5.

2.6.2 Canonical Ensemble

In order to study systems that are closer to those found in nature, it is necessary to consider systems that interact with their surroundings and are not isolated. Under these conditions, the energy is not constant. The canonical ensemble, which facilitates the study of thermally interacting systems, was developed by Gibbs [15]. One way to envision how it is possible to maintain the temperature of the systems in the ensemble is to imagine that an infinite heat sink is in

contact with each system. Thus, any heat generated in any individual system is absorbed by the heat sink and the temperature of the ensemble remains constant.

In the canonical ensemble, the probability distribution function is defined as

$$f(\Gamma) = \frac{e^{-\beta \mathcal{H}_0}}{\int d\Gamma \, e^{-\beta \mathcal{H}_0}} \tag{2.25}$$

where $\beta = \frac{1}{k_B T}$, k_B is Boltzmann's constant and the system is at constant N, V and T.

2.7 Further Considerations

Here we examine some practical considerations for MD simulations: ensemble averaging, phase-space averaging, surface effects and the use of reduced units in simulations.

2.7.1 Ensemble Averaging

In order to determine the macroscopic properties of a system by calculating these properties for individual microstates, it is necessary to average each property over the entire ensemble (§2.4) of microstates. This *ensemble average* is computed as

$$< A(t) >= \frac{\int d\Gamma f(\Gamma, t) A(\Gamma)}{\int d\Gamma f(\Gamma, t)} \tag{2.26}$$

where $A(\Gamma)$ is any phase function and $f(\Gamma, t)$ is the phase-space distribution function(§2.5.1) at time t.

2.7.2 $\mathcal{O}(\frac{1}{N})$ Corrections

Due to constraints in the equations of motion, the number of degrees of freedom in a system are often reduced. For instance, if $\sum_i^N \mathbf{p}_i = 0$ then the number of degrees of freedom will be reduced by d, the Cartesian dimension of the system in question. Further constraints may be commonly encountered in the form of temperature or energy constraints. It is only important to consider these

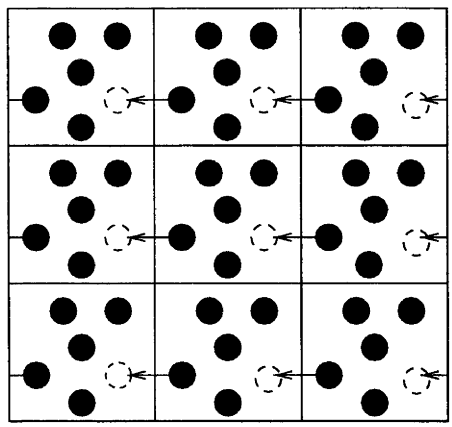


Figure 2.6: Periodic boundary conditions consist of an infinite array of boxes (in two dimensions) or cubes (in three dimensions). As a particle moves to leave a box or cube, its image from a neighbouring box moves to enter from the opposite side or face.

reductions in the number of degrees of freedom when the number of particles in the system are small. If there are many particles in the system, the effect of applying $\mathcal{O}(\frac{1}{N})$ corrections will be small and often unnecessary.

As an example, consider a system of N = 8 particles in two Cartesian dimensions. Normally, this system would have 4N degrees of freedom. Due to momenta constraints, the number of degrees of freedom will be reduced to 4N − 2. If a Gaussian thermostat (§2.9.2) is applied to the system, then the kinetic energy constraint will mean that the number of degrees of freedom of the system in momentum space is 2N − 3.

2.7.3 Periodic Boundary Conditions

Unless one is interested in studying interfaces (liquid-gas/liquid-solid), molecular dynamics simulations that are applied to liquid-state systems are modelled so as to avoid surface effects influencing the results. These effects result from some particles being near a simulated surface of a liquid, while other particles interact well away from this surface. Periodic boundary conditions reduce the effects of using finite N. This technique involves the replication of the simulated unit cell that holds the particles. This box is replicated in all dimensions. Since all repeated cells are identical, we need only consider dynamics in one cell. As a particle moves to leave the primitive cell, its image moves to enter the same

cell from the opposite side and enters the primitive cell with the same velocity as the exiting particle. A diagram illustrating periodic boundary conditions is shown in Figure 2.6.

2.7.4 Reduced Units

In order to simplify MD simulations, reduced units are often employed. Reduced units entail the redefinition of the units for calculating physical quantities from the macroscopic SI system to microscopic units—typically m, σ and ϵ of the Lennard-Jones or WCA interaction potentials discussed in (§2.3). Often, when using reduced units in MD simulations, $\frac{\epsilon}{k_B}$,σ and m are all set to unity [10–12]. Using the definitions of these three units, one may derive the values of all other properties of the system, e.g. the unit of time in reduced units is $\sigma\sqrt{\frac{m}{\epsilon}}$. It is possible to perform MD simulations in reduced units and then convert the calculated macroscopic properties from reduced units to "real" units. When using the LJ potential (§2.3), liquid argon can be reasonably accurately simulated if $\frac{\epsilon}{k_B} = 119.8\text{K}$, $\sigma = 3.405\text{Å}$ and molar mass, $\mathcal{M} = 0.03994\text{kg/mol}$ [10–12]. Reduced units will be used throughout subsequent chapters of this thesis.

2.8 Non-equilibrium algorithms

Non-equilibrium systems are of much interest in science and engineering, as most systems in nature are not at equilibrium. The computer simulation of non-equilibrium systems is largely based on computer algorithms that have been devised during the last 30 years. Many contemporary techniques [10, 11] may be used to solve non-equilibrium problems computationally that would otherwise be intractable in an experimental laboratory. For brevity, only two popular techniques are presented in this section.

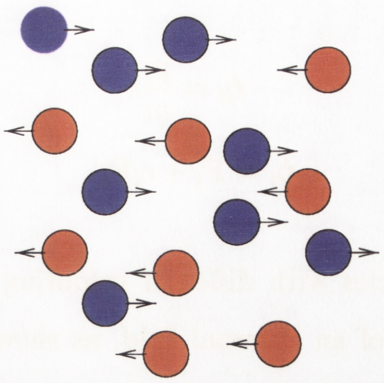


Figure 2.7: Illustration of colour diffusion. The particles in a system are divided into two separate groups which respond to the external field in a different manner. In this example, their response to the external field forces the particles to move in opposite directions. The particles still interact with one another according to the interaction potential defined for the system.

2.8.1 Colour Diffusion

In order to simulate the effects of an external field on a system, it is necessary to add extra terms to the equations of motion, as discussed in (§2.4.1). In those equations of motion, the external field is coupled to the system via two coupling terms, $\mathbf{C}_i$ and $\mathbf{D}_i$. In colour diffusion, the term coupling the external field to the generalised coordinates, $\mathbf{C}_i$ is zero for each particle. The coupling term for the generalised momenta, $\mathbf{D}_i$ is equal to $c_i\mathbf{1}$. The coupling constant c_i is used to mathematically "colour" the particles in the system. If it is desired to divide the particles in a system into two groups, as depicted in Figure 2.7, then $c_i = (-1)^i$. Other schemes for "colouring" the particles in a system may be desired or required, depending on what is under investigation. An example of a different colouring scheme is discussed in Chapter 7.

The colour diffusion Hamiltonian is defined by

$$\mathcal{H}_{col} = \mathcal{H}_0 - \sum_{i=1}^{N} c_i \mathbf{q}_i \cdot \mathbf{F}_e \tag{2.27}$$

where $\mathcal{H}_0$ is the equilibrium Hamiltonian and $\mathbf{F}_e$ is the external colour field. The equations of motion for coupling of the external field to the momenta only are

$$\dot{\mathbf{q}}_i = \frac{\mathbf{p}_i}{m}$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i + c_i \mathbf{F}_e. \tag{2.28}$$

This forces the particles with different colouring to move in opposite directions in the presence of an external field, as shown in Figure 2.7. As the particles pass by one another, they also interact according to the defined interaction potential.

If a field, $\mathbf{F}_e$, is applied in the x-direction only, the colour current density, J_x, for this system is

$$\mathrm{J}_x = \frac{1}{\mathrm{V}} \sum_{i=1}^{\mathrm{N}} c_i p_{xi} \tag{2.29}$$

and the self-diffusion coefficient is then [10]

$$\mathsf{D} = \frac{1}{\beta\rho} \lim_{t\to\infty} \lim_{\mathrm{F}_e\to 0} \frac{\left\langle \mathrm{J}_x(t) \right\rangle}{\mathrm{F}_e}. \tag{2.30}$$

2.8.2 The SLLOD Algorithm

The SLLOD algorithm is another algorithm used for simulating particle motion in the presence of an external field. This type of algorithm can be used to model any type of homogeneous flow. For shear flow, different layers of particles move with different streaming velocities, as illustrated in Figure 2.8. This is a graphical depiction of the "sliding-brick" Lees-Edwards periodic boundary conditions [16]. Other types of flow can be modelled with SLLOD, provided suitable boundary conditions can be devised. Note that the velocity in the x direction is related to the y coordinate of the given imaged cell. The equations of motion for the SLLOD algorithm are similar to those of the DOLLS algorithm [17]. The DOLLS equations of motion are derivable from the DOLLS tensor Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 + \sum_{i=1}^{\mathrm{N}} \mathbf{q}_i \mathbf{p}_i : (\nabla \mathbf{u}(t))^{\mathrm{T}}. \tag{2.31}$$

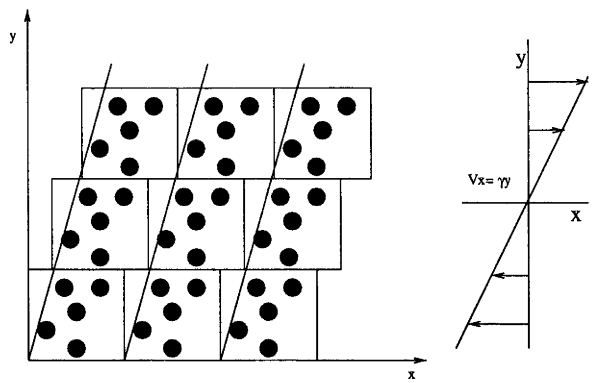


Figure 2.8: Illustration of shear flow in a replicated infinite lattice of boxes (two dimensions). The velocity of the boxes in the x direction is dependent on their coordinates relative to the y axis: $v_x = \gamma y$ where γ is the shear rate.

The equations of motion obtained from this Hamiltonian are

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} + \mathbf{q}_i \cdot \boldsymbol{\nabla}\mathbf{u} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \boldsymbol{\nabla}\mathbf{u} \cdot \mathbf{p}_i \end{aligned} \tag{2.32}$$

where $\boldsymbol{\nabla}\mathbf{u}$ is the velocity gradient tensor.† For planar Couette flow,

$$\boldsymbol{\nabla}\mathbf{u} = \begin{pmatrix} 0 & 0 & 0 \\ \gamma & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.33}$$

where γ is the y-gradient of the x-streaming velocity and is referred to as the strain rate. It is defined by $\gamma \equiv \frac{\partial u_x}{\partial y}$.

The SLLOD equations of motion have superseded those of DOLLS. DOLLS and SLLOD give the same results in the linear regime, but DOLLS gives incorrect results at quadratic or greater order shear rates. As might be expected from the names of these two algorithms, they are closely related. The SLLOD equations of motion are [10]

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} + \mathbf{q}_i \cdot \boldsymbol{\nabla}\mathbf{u} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \mathbf{p}_i \cdot \boldsymbol{\nabla}\mathbf{u}. \end{aligned} \tag{2.34}$$

†The velocity gradient tensor is often referred to as the strain-rate tensor.

Notice that the only difference between the DOLLS and SLLOD equations of motion is the order of the coupling of the momenta to the strain-rate tensor. This difference is enough, however, to prevent the SLLOD equations of motion from being derivable from a Hamiltonian. Even though this is not possible, the phase space compression factor (§2.5.1) for dynamics obtained via the use of the SLLOD equations of motion is zero. In the absence of a thermostat and if the phase-space compression factor is zero, adiabatic incompressibility of phase space (AIΓ) is satisfied. Such is the case with the SLLOD equations of motion.

It is possible to write the SLLOD equations of motion for Couette flow (Equation (2.34)) in an alternate form. If we note that

$$[x \quad y \quad z] \cdot \begin{pmatrix} 0 & 0 & 0 \\ \gamma & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \mathbf{i}\gamma y \tag{2.35}$$

then the SLLOD equations of motion become

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} + \mathbf{i}\gamma y_i \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \mathbf{i}\gamma p_{yi}. \end{aligned} \tag{2.36}$$

The dissipative flux is

$$\dot{\mathcal{H}}_0^{ad} \equiv -\mathsf{P}_{xy}\gamma \mathrm{V}, \tag{2.37}$$

where

$$\mathsf{P}_{xy}\mathrm{V} = \left\langle \sum_{i=1}^{\mathrm{N}} \frac{p_{xi}p_{yi}}{m} - \frac{1}{2}\sum_{i,j}^{\mathrm{N}} x_{ij}\mathrm{F}_{\mathrm{yij}} \right\rangle \tag{2.38}$$

and P_{xy} is the xy-element of the pressure tensor. The shear viscosity, η, is then defined by

$$\eta(\gamma) = \lim_{t\to\infty} -\frac{< \mathsf{P}_{xy}(t) >}{\gamma}. \tag{2.39}$$

2.9 Thermostatting mechanisms

2.9.1 Temperature

The thermodynamic definition of temperature is

$$\frac{1}{\mathrm{T}_t} \equiv \left.\frac{\partial \mathcal{S}}{\partial \mathsf{E}}\right|_{\mathrm{V}}. \tag{2.40}$$

Here $\mathcal{S}$ is the entropy of the system and E the internal energy. There are numerous ways that $\frac{\partial \mathcal{S}}{\partial \mathsf{E}}$ can be evaluated [18,19]. One simple method yields

$$\frac{1}{2} d(\mathrm{N}-1)\mathrm{k_B}\mathrm{T}_k = \frac{1}{2m}\sum_{i=1}^{\mathrm{N}} p_i^2, \tag{2.41}$$

where T_k is the kinetic temperature, d the dimension of the system, N the number of particles, m the mass and p the momentum of each particle, respectively. There are, however, other expressions for the derivative shown in Equation (2.40) (see [19,20]). At equilibrium, the ensemble average of each is identical and gives the thermodynamic temperature.

Away from equilibrium, we will usually take Equation (2.41) to be the temperature. In this thesis, the kinetic temperature is used in all molecular dynamics simulations. When discussing the methodology of MD simulations throughout this thesis, the kinetic temperature is denoted by "T."

In order to keep an MD simulation at constant temperature, it is necessary to implement a thermostatting mechanism. Several approaches for performing constant temperature MD simulations have been developed. Of these methods, two will be discussed that are particularly useful in NEMD simulations.

2.9.2 Gaussian Thermostat

To maintain a system at constant temperature, one approach is to constrain the peculiar kinetic energy to a constant value. One such constraint method is the Gaussian thermostat. The Gaussian thermostatting mechanism involves the introduction of a "friction" term in the equations of motion. This term is

a Lagrange multiplier that has the effect of adding or removing heat from the system to keep the total kinetic energy, K, fixed. If a thermostat is included in the equations of motion (Equation (2.7)), the equations become

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \alpha \mathbf{p}_i. \end{aligned} \tag{2.42}$$

Here α is the Gaussian thermostat multiplier. Since the kinetic energy is a constant of the motion, we have $\mathsf{K} = \frac{1}{2}\frac{\sum_i \mathbf{p}_i \cdot \mathbf{p}_i}{m} = \mathsf{K}_0$, i.e. $\mathbf{p} \cdot \dot{\mathbf{p}} = 0$, so α is defined by

$$\alpha = \frac{\sum_i \mathbf{F}_i \cdot \mathbf{p}_i}{\sum_i \mathbf{p}_i \cdot \mathbf{p}_i}. \tag{2.43}$$

As the kinetic energy of the system will be constant when a Gaussian thermostat is present, an ensemble of such systems is denoted as the isokinetic ensemble. By adding a kinetic energy constraint to the equations of motion, energy is no longer conserved. The differential nature of the Gaussian thermostat facilitates the constraint of a system's temperature instantaneously, i.e. at any moment during an MD simulation, the kinetic temperature of the simulation will be held constant. Note that these equations of motion (Equation (2.42)) cannot be derived from a Hamiltonian (i.e. are non-Hamiltonian).

At equilibrium the equilibrium distribution function is

$$\frac{\mathrm{e}^{-\beta\Phi(q)}\delta(\mathsf{K}(\mathbf{p}) - \mathsf{K}_0)}{\int \mathrm{d}\mathbf{\Gamma}\, \mathrm{e}^{-\beta\Phi(q)}\delta(\mathsf{K}(\mathbf{p}) - \mathsf{K}_0)}, \tag{2.44}$$

where $\frac{3}{2}\mathrm{N}\beta^{-1} = \mathsf{K}_0$. This is canonical in $\mathbf{q}$.

If a Gaussian thermostat is used in a system that is subjected to an external colour field (§2.8.1), then the equations of motion will have terms for both the thermostat and the coupling of the field to the momenta. The equations of motion for such a system are

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}}{m} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i + \mathbf{i}c_i\mathrm{F}_e - \alpha\Big(\mathbf{p}_i - \mathbf{i}c_i\frac{\mathrm{J}_x}{\rho}\Big). \end{aligned} \tag{2.45}$$

Here ρ is the number density of the system and α, the Gaussian thermostat multiplier has the form

$$\alpha = \frac{\sum \mathbf{F}_i \cdot \left(\mathbf{p}_i - \mathbf{i}c_i \frac{\mathrm{J}_x}{\rho}\right)}{\sum \mathbf{p}_i \cdot \left(\mathbf{p}_i - \mathbf{i}c_i \frac{\mathrm{J}_x}{\rho}\right)} \tag{2.46}$$

The velocities we calculate in our equations of motion under these conditions are the peculiar velocities of each particle relative to the streaming velocity of each "coloured" species.

Likewise, for a system using the SLLOD equations of motion (Equation (2.36)), the introduction of a thermostat will add an extra term to the equations of motion for the momenta

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} + \mathbf{i}\gamma y_i \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \mathbf{i}\gamma p_{yi} - \alpha \mathbf{p}_i. \end{aligned} \tag{2.47}$$

The form of the Gaussian thermostat multiplier for the SLLOD equations of motion is

$$\alpha = \frac{\sum \left(\mathbf{F}_i \cdot \mathbf{p}_i - \gamma p_{xi} p_{yi}\right)}{\sum p_i^2}. \tag{2.48}$$

2.9.3 Nosé-Hoover Thermostat

The equations of motion for a Nosé-Hoover thermostated equilibrium system are

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \zeta(\Gamma)\mathbf{p}_i \\ \dot{\zeta} &= \frac{1}{Q}\left[\sum \frac{\mathbf{p}_i^2}{m} - (g+1)\mathrm{k_B}T\right]. \end{aligned} \tag{2.49}$$

Here Q is the effective mass of a heat bath and g is the number of degrees of freedom in the system.

The Nosé-Hoover (NH) thermostatting mechanism [21,22] is similar to the

Gaussian Thermostat (§2.9.2), in that it provides a mechanism for restricting the temperature to a certain value. There are several distinct differences between the two, however. While the Gaussian thermostat has a differential feedback mechanism, the Nosé-Hoover thermostat involves the integration of the Lagrangian thermostat multiplier at the same time the equations of motion are integrated.

The NH thermostat also differs from the Gaussian thermostat in that it generates the canonical ensemble. That is, if one thermostats a system using the NH mechanism, upon equilibration the system will have a canonical distribution (Equation (2.25)). Furthermore, the integral nature of the NH thermostat results in a system temperature that is constant *on average.* Instantaneous temperature values calculated during the course of an MD simulation utilising an NH thermostat will not generally be equal to the desired constant temperature value for the system.

At equilibrium the Nosé-Hoover equilibrium distribution function is

$$f_{\mathrm{NH}}(\mathbf{\Gamma}) = \frac{\mathrm{e}^{-\beta(\mathcal{H}_0+\frac{1}{2}Q\zeta^2)}}{\int \mathrm{d}\mathbf{\Gamma}\,\mathrm{d}\zeta\,\mathrm{e}^{-\beta(\mathcal{H}_0+\frac{1}{2}Q\zeta^2)}} \tag{2.50}$$

i.e. it is the usual canonical ensemble in $\mathbf{\Gamma}$.

Chapter 3

The Fluctuation Theorem

3.1 Introduction

In the previous chapter tools were discussed that are necessary for developing molecular dynamics simulations which model microscopic states and compute instantaneous values of microscopic phase functions. In this chapter, we will use these tools to examine the nature of the relationship between the Fluctuation Theorem and the Second Law of Thermodynamics.

The Second Law is a general statement that is fundamental to statistical thermodynamics. However, as will be discussed in (§3.2), it is only strictly applicable to macroscopic systems. It is possible that by observing the microscopic properties of a system, one could witness events that occur in a manner opposite to what is predicted by the Second Law.

The Fluctuation Theorem (FT) is a mathematical expression that gives the probability that the Second Law of Thermodynamics may be violated for a finite system observed for a finite time. The Fluctuation Theorem is therefore an expression that is applicable to the microstates of a system and is simply a restatement of the Second Law for infinite thermostatted systems observed for infinite times.

The FT was first discovered by Evans, Cohen and Morriss in 1993 [23]. In this original paper, they showed that for a system in the microcanonical ensemble, the probability that the time-averaged irreversible entropy production,

$\overline{\Sigma}_t = \frac{1}{t}\int_0^t \mathrm{d}s\ \Sigma(s)$, takes on a positive value increases exponentially with time and system size, since Σ is extensive.

$$\lim_{t\to\infty} \ln\left[\frac{\Pr(\frac{\overline{\Sigma}_t}{k_B} = \mathrm{A})}{\Pr(\frac{\overline{\Sigma}_t}{k_B} = -\mathrm{A})}\right] = \mathrm{A}t \tag{3.1}$$

This is exactly what one would expect from the Second Law of Thermodynamics. The FT is important for several reasons. It is one of the few mathematical expressions in non-equilibrium statistical mechanics that is valid even far from equilibrium. The FT is an exact mathematical expression—no assumptions or approximations need be made in its application. Green-Kubo relations may be also derived from the application of the FT (§3.10) [24].

3.2 Second Law of Thermodynamics

Mechanics is time reversible. If one were to make a movie of the motion of a system of particles, this movie, when played backwards, would still be a solution of the laws of mechanics. An example of one such process is a ball at rest on the floor that spontaneously begins to bounce due to energy it absorbs from the floor and surroundings in the form of heat. The formulation of the Second Law of Thermodynamics is a direct statement of the fact that such events are never observed.

The Second Law of Thermodynamics was first proposed by Rudolf Clausius in 1850 [25]. At this time, Clausius stated

> "Es ist unmöglich, Wärme von einem kälteren zu einem wärmeren Körper ohne Compensation zu überführen." [26]

The Second Law has been stated in several other ways, but one of its best-known definitions is that given by William Thomson (a.k.a. Lord Kelvin) [27]

> "It is impossible, by means of inanimate material agency, to derive mechanical effect from any portion of matter by cooling it below the temperature of the coldest surrounding objects."

These statements are central to thermodynamics and allow for the definition of entropy, a state function. The concept of entropy was first introduced by Clausius and the statistical definition of entropy was proposed by Boltzmann and later written by Planck [28] as

$$\mathcal{S} = \mathrm{k}\ \mathrm{logW}$$

where, in Planck's notation, k is Boltzmann's constant, W is the weight of the most probable configuration of the system and the logarithm is base e. In terms of the entropy, the Second Law may be restated [29]

> The entropy of an isolated system increases in the course of a spontaneous change: $\Delta\mathcal{S}_{tot} > 0$

where $\mathcal{S}_{tot}$ is the total entropy of an isolated system. Entropy production for spontaneous processes is irreversible. This is a statement that is applicable only to macroscopic systems.

The Second Law is important not only in statistical mechanics, but is central to other areas of science as well. It is important for things as diverse as understanding the nature of evolution [30] and the thermodynamics of entities as exotic as black holes [31]. Recently, a quantum derivation of the Second Law has been proposed [32].

3.3 The FT and the Second Law

Questions as to the basis of a causal, as opposed to an anti-causal, universe have been debated for more than a century. Fundamental questions such as *Why are there some events that never happen in a reverse direction?* have plagued philosophy and science during this time, and statistical thermodynamics in particular [33]. Examples of such events include a waterfall flowing upwards or a jet engine that runs in reverse, taking in exhaust fumes and heat to generate oxygen and kerosene.

Josef Loschmidt, a mathematician and contemporary colleague of Boltzmann, argued that since mechanics is reversible, the Second Law of Thermody-

namics cannot be proven from mechanics.* Boltzmann dubbed this argument **Loschmidt's Paradox** [34]. As a result of Loschmidt's continuous arguments and examples of mathematical reversibility, Boltzmann conceded

> "To be sure, as soon as one looks at bodies of such small dimension that they contain only a very few molecules, the validity of this theorem [the Second Law] must cease." [28]

However, experience shows that the probability of such events taking place must be extremely small. Consideration of Loschmidt's paradox leads directly to the Fluctuation Theorem. These negative *fluctuations* in the entropy predicted by Loschmidt happen on a microscopic scale.

While it was Evans, Cohen and Morriss who first proposed the Fluctuation Theorem, and Evans and Searles who gave the first mathematical derivation of the theorem, it was Maxwell and Loschmidt who first realised that the Second Law of Thermodynamics is not absolute. [34,35] Loschmidt's constant counter-arguments to Boltzmann's early theories on what later became the Second Law prompted Boltzmann to refine some of his beliefs as to the statistical mechanical nature of the universe. [28] The argument of mechanical reversibility and its relationship to the Second Law of Thermodynamics is still a topic to which much contemporary scientific research has been applied; see, for example, [36–40].

3.4 The Fluctuation Theorem

Since its discovery, several versions of the FT have been studied. The application of the Fluctuation Theorem to a series of finite segments along a single, steady-state phase space trajectory has been called a *dynamical* FT (DFT). This is the version of the Fluctuation Theorem that was examined in the microcanonical ensemble in the original paper by Evans, Cohen and Morriss in 1993. In 1994, Evans and Searles derived a version of the FT [41] that

*Although James Clerk Maxwell first examined the possibility of mechanical reversibility with respect to the Second Law, it was Loschmidt who first published such arguments [28].

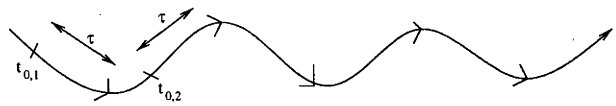


Figure 3.1: Illustration of the dynamical FT. Sub-segments of a single trajectory segment are defined. Phase properties are averaged over each of these sub-segments, yielding an ensemble average.

used non-equilibrium segments spawned from an equilibrium trajectory. This *transient* version of the FT (TFT) was first examined in the microcanonical ensemble. Subsequently, Gallavotti and Cohen clarified the proof of the DFT in the microcanonical ensemble using the Sinai-Ruelle-Bowen measure [42]. A third version of the FT was proposed by Evans and Searles in 1994 [41]. Originally also in the microcanonical ensemble, this version of the FT involves non-equilibrium trajectory segments spawned from an equilibrium trajectory, as in the TFT. However, the non-equilibrium trajectory segments are allowed to evolve until they have reached the steady state and the ensemble averaging is done only over the steady-state part of the trajectory. This version of the FT is therefore denoted the *steady-state* FT (SSFT). Later [43], the TFT and SSFT were generalised to several other ensembles. A more detailed description of each of these versions of the FT is given later in this section.

3.4.1 Dynamical Fluctuation Theorem

In the dynamical version of the FT, a single trajectory is propagated through phase space. This trajectory is partitioned into smaller sub-trajectories of length τ that start at regular points along the trajectory, e.g. $t_{0,1}$, $t_{0,2}$, etc. Figure 3.1 depicts this process graphically.

The DFT is only valid in the limit $t \to \infty$ and assumes that the system in question is ergodic. That is, its ensemble average is equal to its time average for all $\mathbf{\Gamma}(0)$. The premise behind the DFT is that if one observes the instantaneous values of a given phase function for a finite time (a sub-segment of duration $t = \tau$), occasionally fluctuations in the direction opposite to the normal response for this phase function will be seen. If, for example, one monitors the value of the entropy production in a molecular dynamics simulation for a system with

a small number of particles and for a finite time, one will see that if the system is sufficiently close to equilibrium, the entropy production will sometimes be negative. From the mathematical form of the Fluctuation Theorem (Equation (3.1)), we see that the possibility of observing these negative values for the entropy decreases with time. Note that in order for the DFT to hold, the phase function must be averaged over sub-segments of *equal* length. It is also important to note that at equilibrium, the probability of observing a phase value and a phase anti-value is assumed to be equal. This assumption is based on the relation $\mathcal{H}(p,q) = \mathcal{H}(-p,q)$.

3.4.2 Transient Fluctuation Theorem

In the TFT, an equilibrium trajectory is propagated through phase space. Periodically, a non-equilibrium side trajectory is spawned from this trajectory and the entropy production is averaged over the length of this side, or transient, trajectory (see Figure 3.2). A collection of many of these transient trajectories comprises an *ensemble* (§2.4). The average of the individual values of a phase function over the length of the trajectory is the time average

$$\overline{\Psi}_t = \frac{1}{t}\int_0^t \mathrm{d}s\, \Psi(s) \tag{3.2}$$

where Ψ is any phase function. The ensemble average is then

$$< \overline{\Psi}_t >= \frac{\sum_{j=1}^{\mathrm{N}} \overline{\Psi}_{tj}}{\sum_{j=1}^{\mathrm{N}} 1}, \tag{3.3}$$

where N is the total number of transient trajectories. It is therefore imperative that the length of each of the transient trajectories be the same to facilitate this averaging. A graphical depiction of this process is shown in Figure 3.2. Note that for the TFT, ergodicity is allowed, but not required.

It has been shown that the TFT is valid in various ensembles including the microcanonical [24,41] and the canonical [43] ensembles. In both of these cases, it is necessary to apply an external field to force the system away from equilibrium and generate transient trajectories. In terms of MD simulations, it is

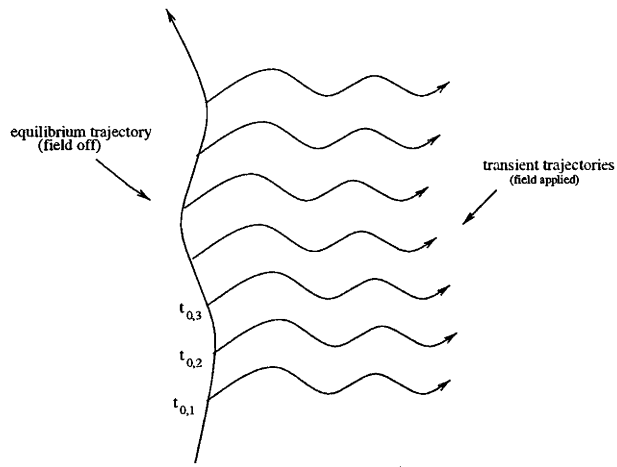


Figure 3.2: Illustration of the transient fluctuation theorem. Non-equilibrium (transient) trajectories are spawned from a single equilibrium trajectory. Phase properties are averaged over the length of each transient trajectory and ensemble averages are calculated from the average over all transient trajectories.

necessary to propagate an equilibrium backbone trajectory via the use of Newton's equations (microcanonical case) or thermostat (canonical or isokinetic case). If an equilibrium constant temperature backbone trajectory is created, the resulting ensemble will be dependent on what type of thermostat is used. If a Gaussian thermostat (§2.9.2) is used, the equilibrium ensemble will be isokinetic, while if a Nosé-Hoover thermostat is used, it will be canonical. The TFT can be extended to the SSFT if the transient trajectories are allowed to propagate for a sufficiently long period of time.

3.4.3 Steady-state Fluctuation Theorem

The SSFT is similar to the TFT in that an equilibrium backbone trajectory is propagated through Γ-space from which non-equilibrium transient trajectories are spawned. In the SSFT, however, initial transient segments are ignored and the entropy production is averaged only over the steady-state portion of the ensemble of non-equilibrium trajectories. The amount of time needed to reach a steady state varies depending on the nature of the system.

One may verify that the system has reached a steady state by checking that the ensemble averaged values of phase functions are time independent. Figure 3.3 shows an example of a SSFT trajectory, while in Figure 3.4, the

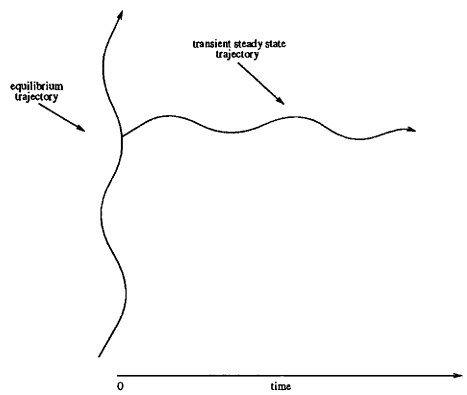


Figure 3.3: A graphical depiction of the evolution of a SSFT simulation. If the transient trajectory is propagated for a sufficiently long period of time, the steady state will eventually be reached.

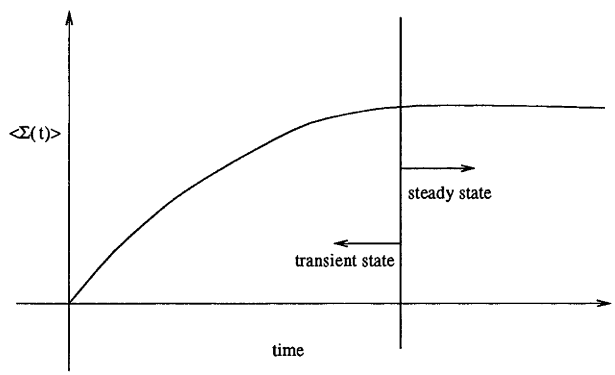


Figure 3.4: An illustration of the difference between the TFT and SSFT. The TFT involves only transient trajectory segments, while the SSFT uses those same segments, but which are propagated for a longer time to reach a point where the entropy production is in the steady-state. In the figure, $< \Sigma(t) >$ is the ensemble average of the rate of entropy production.

corresponding entropy production is shown. One can see that there is initially a time where the entropy production of the non-equilibrium trajectory has not yet reached the steady state (called the transient state). As the trajectory is propagated for a longer period of time, however, the steady state is eventually achieved. Figure 3.4 also gives a graphical representation of the difference between the transient and steady-state version of the fluctuation theorem.

3.5 Fluctuation Theorem Derivation

Chapter 2 defined basic concepts which are vital for the derivation of the Fluctuation Theorem. In this section, the concepts developed there are applied

to facilitate the derivation. In (§3.5.1) through (§3.5.4) we discuss how each of the fundamental concepts are important in deriving the FT. Derivation of the FT involves considering sets of trajectories and their unique conjugate trajectories, whose initial points are sampled from a known distribution (e.g. microcanonical) and hence whose probability of being observed is known. Below we consider how we calculate these probabilities and how the conjugate trajectories are constructed.

3.5.1 Γ-space Distribution

If one examines the probability that a trajectory segment will be found within a given volume in phase space, it is necessary to consider the phase-space distribution function for a given volume element at some point $\mathbf{\Gamma}(t)$ in phase space at an arbitrary time t. Equation (2.16) may be expressed in terms of the probability of finding in a given phase volume about a phase point $\mathbf{\Gamma}(t)$

$$\Pr(\delta V(\mathbf{\Gamma}(t),t)) = f(\mathbf{\Gamma}(t),t)\delta V(\mathbf{\Gamma}(t),t) \tag{3.4}$$

Here $f(\mathbf{\Gamma}(t),t)$ is the phase-space distribution function and δV is an infinitesimal phase space volume as defined in Equation (2.15).

3.5.2 Time-reversal Mapping

Consider a system that has reversible equations of motion. Using the Liouville operator [10] associated with the time derivative of a phase variable† we can move from an initial point in phase space, $\mathbf{\Gamma}(0)$, to some later point, $\mathbf{\Gamma}(t)$

$$\mathbf{\Gamma}(t) = \mathrm{e}^{iL(\mathbf{\Gamma})t}\mathbf{\Gamma}(0). \tag{3.5}$$

Using this operator and the reversibility of the equations of motion, we can construct a unique conjugate trajectory for any selected trajectory (see Figure 3.5). We start at a point $\mathbf{\Gamma}(0)$ and propagate in the forward direction for a

†In contrast to the distribution function Liouville operator (Equation 2.22).

time $\frac{t}{2}$. Since the equations of motion are time-reversible, we can apply the time-reversal mapping, whereby $(q,p) \to (q,-p)$. So

$$\mathbf{\Gamma}(\frac{t}{2}) \to \mathbf{\Gamma}^*(\frac{t}{2}) = \mathrm{M}^{\mathrm{T}}(\mathbf{\Gamma}(\frac{t}{2})). \tag{3.6}$$

If the equations of motion are then allowed to iterate in the forward direction, we will reach $\mathbf{\Gamma}^*(t)$ at time t, which is related to $\mathbf{\Gamma}(0)$ via M^{T}. We can see how the time-reversal mapping and propagation of the trajectory in the forward direction yield a time-reversed system

$$\begin{aligned}
\mathbf{\Gamma} &= \mathrm{e}^{-i\mathrm{L}(\mathbf{\Gamma})t}\mathrm{e}^{i\mathrm{L}(\mathbf{\Gamma})t}\mathbf{\Gamma} \\
&= \mathrm{M}^{\mathrm{T}}\mathrm{M}^{\mathrm{T}}\mathrm{e}^{-i\mathrm{L}(\mathbf{\Gamma})t}\mathrm{e}^{i\mathrm{L}t}\mathbf{\Gamma} \\
&= \mathrm{M}^{\mathrm{T}}\mathrm{e}^{\mathrm{M}^{\mathrm{T}}[-i\mathrm{L}(\mathbf{\Gamma})]t}\mathrm{e}^{\mathrm{M}^{\mathrm{T}}[i\mathrm{L}(\mathbf{\Gamma})]t}\mathbf{\Gamma} \\
&= \mathrm{M}^{\mathrm{T}}\left(\mathrm{e}^{i\mathrm{L}(\mathbf{\Gamma})t}\right)\mathrm{M}^{\mathrm{T}}\mathrm{e}^{i\mathrm{L}(\mathbf{\Gamma})]t}\mathbf{\Gamma}
\end{aligned} \tag{3.7}$$

which implies

$$\mathrm{M}^{\mathrm{T}}\mathrm{e}^{i\mathrm{L}(\mathbf{\Gamma})t}\mathrm{M}^{\mathrm{T}}\mathrm{e}^{i\mathrm{L}(\mathbf{\Gamma})t} = 1. \tag{3.8}$$

Furthermore, if we start at $\mathbf{\Gamma}^*(\frac{t}{2})$ and propagate backward in time (effectively following $\mathbf{\Gamma}$ from $\frac{t}{2}$ in the forward direction), we will reach $\mathbf{\Gamma}^*(0)$ at time t, that is related to $\mathbf{\Gamma}(t)$ by M^{T}. The trajectories $\mathbf{\Gamma}(0;t)$ and $\mathbf{\Gamma}^*(0;t)$ are related by the time reversal mapping.

3.5.3 Phase Volume and Γ-space Compression

If one gives some thought to arguments of reversibility and the FT, one may ask *Why are initial states that eventually lead to Second Law violating states so rarely observed?* In this subsection, time reversibility and trajectory mappings in relation to phase volumes are examined in detail.

We derive the TFT for a system which initially is microcanonical and which evolves under the dual influence of a dissipative external field and an ergostat. Since derivation of the FT requires the probability of observing infinitesimal

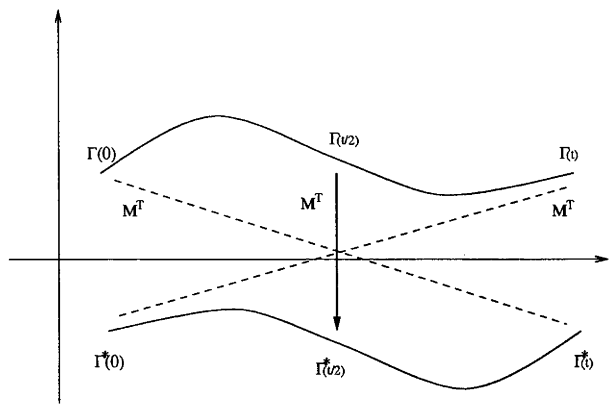


Figure 3.5: A trajectory that has a time-reversal mapping applied to it at time $\tau = \frac{t}{2}$. Starting from $\mathbf{\Gamma}^*(0)$ and propagating the trajectory in the same direction, an anti-trajectory (or anti-causal trajectory) is created. The endpoint of this trajectory ($\mathbf{\Gamma}^*(\frac{t}{2})$) is related to the initial point of the causal trajectory via the time-reversal mapping, M^{T}.

phase volumes near a phase point (see Equation (3.4)), it is necessary to determine the phase volumes about $\mathbf{\Gamma}(0)$ and $\mathbf{\Gamma}^*(0)$, as discussed in (§3.5.2). These can be obtained by considering the change in phase volume under the time reversal mapping and under propagation using the equations of motion.

For the mapping M^{T} shown above, the Jacobian is unity

$$\begin{aligned} \mathrm{J}(\mathrm{M}^{\mathrm{T}}) &= \frac{\partial \mathrm{M}^{\mathrm{T}}(\mathbf{\Gamma})}{\partial \mathbf{\Gamma}} = \frac{\partial \mathrm{M}^{\mathrm{T}}(q,p)}{\partial(q,p)} \\ &= \frac{\partial(q,-p)}{\partial(q,p)} = \begin{vmatrix} \frac{\partial q}{\partial q} & \frac{\partial -p}{\partial q} \\ \frac{\partial q}{\partial p} & \frac{\partial -p}{\partial p} \end{vmatrix} \\ &= \begin{vmatrix} \mathcal{I} & 0 \\ 0 & -\mathcal{I} \end{vmatrix} = 1 \end{aligned} \tag{3.9}$$

and the volume of the system remains the same after the time-reversal mapping is applied. This implies that $\delta V(\mathbf{\Gamma}^*(0)) = \delta V(\mathbf{\Gamma}(t))$. This is illustrated in Figure 3.6

The infinitesimal phase space volume about a phase point will change according to

$$\delta V(\mathbf{\Gamma}(t), t) = \delta V(\mathbf{\Gamma}(0), 0) \mathrm{e}^{\overline{\Lambda}_t t}. \tag{3.10}$$

Here $\overline{\Lambda}_t$ is the time-averaged phase space compression factor and $\mathrm{e}^{\overline{\Lambda}_t t}$ is the

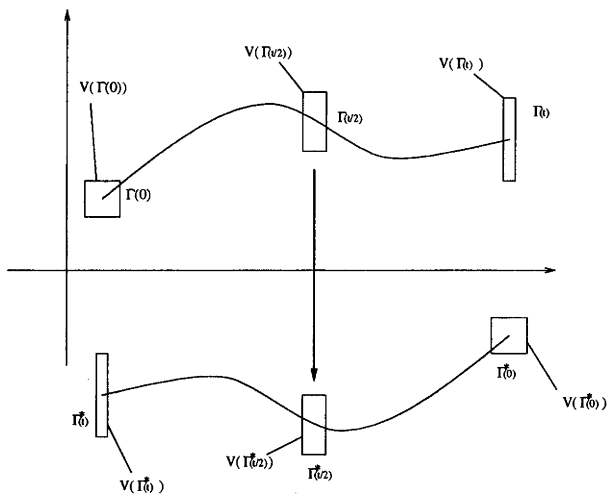


Figure 3.6: The change in phase volume about a trajectory $\mathbf{\Gamma}(0;t)$ and its time-reversed conjugate $\mathbf{\Gamma}^*(0;t)$.

propagator for the phase volume. The mathematical form of $\Lambda(\mathbf{\Gamma})$ is given in (§2.5.1). The time-average of $\Lambda(t)$ is defined by

$$\overline{\Lambda}_t = \frac{1}{t}\int_0^t ds\, \Lambda(\mathbf{\Gamma}(s)) \tag{3.11}$$

If $\Lambda(\mathbf{\Gamma}) = 0$ then $\frac{df}{dt} = 0$, which implies that the density of the phase space volume elements does not change as they move through phase space.

The phase volume elements at the start and end of trajectories and anti-trajectories are related via Equation (3.10) and the fact that the Jacobian of the time-reversal mapping is unity. If we examine an example of the application of the time-reversal mapping with respect to the phase-volume elements of the two trajectories, it can be seen how the phase-volume elements are related (see Figure 3.6).

Consider a trajectory (Figure 3.6) that is propagated through $\mathbf{\Gamma}$-space for some time t. At time $\frac{t}{2}$ a time-reversal mapping described previously (§3.5.2) is applied to the phase point $\mathbf{\Gamma}(\frac{t}{2})$. The mapped phase point is designated as $\mathbf{\Gamma}^*(\frac{t}{2})$ and the conjugate trajectory is constructed as described in (§3.5.2).

Now we can examine the changes of the phase-volume elements about these trajectories. For the forward trajectory, the phase volume at $\frac{t}{2}$ is

$$\delta V(\mathbf{\Gamma}(\frac{t}{2}), \frac{t}{2}) = \delta V(\mathbf{\Gamma}(0), 0) \mathrm{e}^{\int_0^{\frac{t}{2}} \mathrm{d}s\, \Lambda(\mathbf{\Gamma}(s))} \tag{3.12}$$

and that at t is

$$\delta V(\mathbf{\Gamma}(t),t) = \delta V(\mathbf{\Gamma}(0),0)\mathrm{e}^{\int_0^t \mathrm{d}s\, \Lambda(\mathbf{\Gamma}(s))} \tag{3.13}$$

The phase volume after the time reversal mapping at $\frac{t}{2}$ is

$$\delta V(\mathbf{\Gamma}^*(\frac{t}{2}),\frac{t}{2}) = \delta V(\mathbf{\Gamma}(\frac{t}{2}),\frac{t}{2}) \tag{3.14}$$

and the change in phase volume for the conjugate trajectory is

$$\begin{aligned}\delta V(\mathbf{\Gamma}^*(0),0) &= \delta V(\mathbf{\Gamma}^*(\frac{t}{2}),\frac{t}{2})\mathrm{e}^{\int_{\frac{t}{2}}^0 \mathrm{d}s\, \Lambda(\mathbf{\Gamma}^*(s))} \\ &= \delta V(\mathbf{\Gamma}(t),t)\end{aligned} \tag{3.15}$$

and

$$\begin{aligned}\delta V(\mathbf{\Gamma}^*(t),t) &= \delta V(\mathbf{\Gamma}^*(\frac{t}{2}),\frac{t}{2})\mathrm{e}^{\int_{\frac{t}{2}}^t \mathrm{d}s\, \Lambda(\mathbf{\Gamma}^*(s))} \\ &= \delta V(\mathbf{\Gamma}(0),0)\end{aligned} \tag{3.16}$$

Since $\delta V(\mathbf{\Gamma}(t))$ is related to $\delta V(\mathbf{\Gamma}^*(0))$ via the time-reversal mapping M^{T}, and the Jacobian of this mapping is unity, the volumes will be the same. The ratio of these volumes is equal to the exponential integral phase space compression factor of the system

$$\begin{aligned}\frac{\delta V(\mathbf{\Gamma}^*(0),0)}{\delta V(\mathbf{\Gamma}(0),0)} &= \frac{\delta V(\mathbf{\Gamma}(t),t)}{\delta V(\mathbf{\Gamma}(0),0)} \\ &= \mathrm{e}^{\int_0^t \mathrm{d}s\, \Lambda(s)}.\end{aligned} \tag{3.17}$$

From Equation (3.17), we can see

$$\frac{\delta V(\mathbf{\Gamma}^*(t),t)}{\delta V(\mathbf{\Gamma}^*(0),0)} = \mathrm{e}^{\int_0^t \mathrm{d}s\, \Lambda(\mathbf{\Gamma}^*(s))}. \tag{3.18}$$

Using Equation (3.4) and the fact that the phase-space distribution function

for the microcanonical ensemble (§2.6.1) is $f(\mathbf{\Gamma}) = \frac{1}{\int_{E<\mathcal{H}(\mathbf{\Gamma})<E+\Delta E} d\mathbf{\Gamma}}$, we have

$$\ln\left[\frac{\Pr(\delta V(\mathbf{\Gamma}(0),0))}{\Pr(\delta V(\mathbf{\Gamma}^*(0),0))}\right] = -\overline{\Lambda}_t t, \tag{3.19}$$

since the ratio of the probability distribution functions about any two points on the energy hypersurface (e.g. $\mathbf{\Gamma}(0)$ and $\mathbf{\Gamma}(t) = \mathrm{M}^{\mathrm{T}}(\mathbf{\Gamma}^*(0))$) is unity and the natural logarithm is zero. This means that if the time average of the phase-space compression factor for a trajectory $\mathbf{\Gamma}$ is negative, then this trajectory is exponentially more likely to be observed than the conjugate anti-trajectory (which has $\int ds\, \Lambda > 0$).

This then provides an answer to the question posed at the beginning of this subsection. Initial states that evolve to give Second Law violating trajectories are rarely observed due to the small probability of observing their initial phases (they only occupy a small volume in phase space). Phase volumes are therefore fundamental in determining if a system may violate the Second Law.

3.5.4 General Dissipation Function

In the last paragraphs of (§3.5.3) we considered the probability of observing trajectories and their conjugates for a microcanonical system. Here we consider the general situation.

Taking the ratio of probabilities of phase-space volume elements, one arrives at the expression

$$\begin{aligned}\frac{\Pr(\delta V(\mathbf{\Gamma}(0),0))}{\Pr(\delta V(\mathbf{\Gamma}^*(0),0))} &= \frac{f(\mathbf{\Gamma}(0),0)\delta V(\mathbf{\Gamma}(0),0)}{f(\mathbf{\Gamma}^*(0),0)\delta V(\mathbf{\Gamma}^*(0),0)} \\ &= \frac{f(\mathbf{\Gamma}(0),0)}{f(\mathbf{\Gamma}(t),0)} \mathrm{e}^{-\overline{\Lambda}_t t},\end{aligned} \tag{3.20}$$

where we have assumed that the distribution function is even under the time-reversal mapping of the phase point, i.e. $f(\mathrm{M}^{\mathrm{T}}(\mathbf{\Gamma}),0) = f(\mathbf{\Gamma},0)$.[‡] Equation

[‡]This assumption is not necessary, but simplifies the results and is true for equilibrium distribution functions.

(3.20) may be rewritten as

$$\frac{f(\mathbf{\Gamma}(0),0)}{f(\mathbf{\Gamma}(t),0)}\mathrm{e}^{-\overline{\Lambda}_t t} = \mathrm{e}^{\overline{\Omega}_t t}, \tag{3.21}$$

where $\overline{\Omega}_t$ is the time-average of a general dissipation function [43] defined by

$$\overline{\Omega}_t t \equiv \int_0^t \mathrm{d}s\ \Omega(\mathbf{\Gamma}(s)) = \ln\left(\frac{f(\mathbf{\Gamma}(0),0)}{f(\mathbf{\Gamma}(t),0)}\right) - \int_0^t \mathrm{d}s\ \Lambda(\mathbf{\Gamma}(s)). \tag{3.22}$$

Introduction of this general dissipation function simplifies many of the mathematical equations that will follow.

It is now simple to derive the ratio of the probability of observing a particular time-averaged value of a given phase function to its corresponding negative value

$$\begin{aligned}
\frac{\Pr(\overline{\Omega}_t = \mathrm{A})}{\Pr(\overline{\Omega}_t = -\mathrm{A})} &= \frac{\sum_{i|\overline{\Omega}_t=\mathrm{A}} \Pr(\delta V(\mathbf{\Gamma}_i(0),0))}{\sum_{i|\overline{\Omega}_t=-\mathrm{A}} \Pr(\delta V(\mathbf{\Gamma}_i(0),0))} \\
&= \frac{\sum_{i|\overline{\Omega}_t=\mathrm{A}} \Pr(\delta V(\mathbf{\Gamma}_i(0),0))}{\sum_{i|\overline{\Omega}_t=\mathrm{A}} \Pr(\delta V(\mathbf{\Gamma}_i^*(0),0))} \\
&= \frac{\sum_{i|\overline{\Omega}_t=\mathrm{A}} \Pr(\delta V(\mathbf{\Gamma}_i(0),0))}{\sum_{i|\overline{\Omega}_t=\mathrm{A}} \Pr(\delta V(\mathbf{\Gamma}_i(0),0))\frac{f(\mathbf{\Gamma}(t),0)}{f(\mathbf{\Gamma}(0),0)}\mathrm{e}^{\overline{\Lambda}(t)t}} \\
&= \frac{\sum_{i|\overline{\Omega}_t=\mathrm{A}} \Pr(\delta V(\mathbf{\Gamma}_i(0),0))}{\sum_{i|\overline{\Omega}_t=\mathrm{A}} \Pr(\delta V(\mathbf{\Gamma}_i(0),0))\mathrm{e}^{-\overline{\Omega}_t t}} \\
&= \mathrm{e}^{\overline{\Omega}_t t}
\end{aligned}$$

$$\Longrightarrow \ln\left(\frac{\Pr(\overline{\Omega}_t = \mathrm{A})}{\Pr(\overline{\Omega}_t = -\mathrm{A})}\right) = \overline{\Omega}_t t. \tag{3.23}$$

Here the relationship between trajectories and their time-reversed conjugate trajectories has been used to allow us to change the sums of the probability of finding a given trajectory segment in a volume element $\delta V(\mathbf{\Gamma}^*)$ with $\overline{\Lambda}_t = -\mathrm{A}$, to sum over those phase volumes $\delta V(\mathbf{\Gamma})$ associated with a general dissipation function with value A. Equation (3.10) has then been used to obtain the

expression of the ratio in terms of the general dissipation function.

3.6 The FT in Other Ensembles

Although the Fluctuation Theorem has been shown to be valid in several different initial ensembles, the precise form of the FT depends directly on the initial ensemble. This section discusses how the FT differs between the ensembles and completes the derivation for each different case.

3.6.1 FT in the Canonical Ensemble

If one combines Equation (3.22) with the definition of the probability distribution function in the canonical ensemble (Equation 2.25), an expression for the FT in the canonical ensemble is obtained [43],

$$\begin{aligned}\int_0^t \mathrm{ds}\, \Omega(\mathbf{\Gamma}(\mathrm{s})) &= \ln\left(\frac{\mathrm{e}^{-\beta\mathcal{H}(0)}}{\mathrm{e}^{-\beta\mathcal{H}(t)}}\right) - \overline{\Lambda}_t t \\ &= \ln\left(\mathrm{e}^{\beta[\mathcal{H}(t)-\mathcal{H}(0)]}\right) - \overline{\Lambda}_t t \\ &= \beta\int_0^t \dot{\mathcal{H}}(s)\,\mathrm{d}s - \overline{\Lambda}_t t\end{aligned} \tag{3.24}$$

The form of the phase-space compression factor, Λ, will depend on the thermostatting mechanism present for the system. If no thermostat is present and the equations of motion are Hamiltonian, then from the Liouville equation (Equation (2.21)), we have $\Lambda(\mathbf{\Gamma}) \equiv 0$. The form of Λ in the presence of a Gaussian thermostat is discussed in the next section. The form of $\dot{\mathcal{H}}$ is dependent on the dynamics of the system. As different dynamics are discussed later in this thesis, the form of $\dot{\mathcal{H}}$ will be shown in each case.

3.6.2 FT in the Isokinetic Ensemble

The FT in the isokinetic ensemble is much like that in the canonical ensemble. In the isokinetic ensemble, the probability distribution function is

$$f(\mathbf{\Gamma}) = \mathrm{e}^{-\beta\mathcal{H}(\mathbf{\Gamma})}\delta(\mathsf{K}(\mathbf{\Gamma}) - \mathsf{K}_0) \tag{3.25}$$

where $\mathsf{K}(\mathbf{\Gamma})$ is the kinetic energy at $\mathbf{\Gamma}$ and K_0 is the value to which the kinetic energy is constrained.[§]

A combination of Equations (3.25) and (3.22) yields [43]

$$\begin{aligned}\int_0^t \mathrm{ds}\ \Omega(\mathbf{\Gamma}(\mathrm{s})) &= \ln\left(\frac{\mathrm{e}^{-\beta\mathcal{H}(0)}\delta(\mathsf{K}(0) - \mathsf{K}_0)}{\mathrm{e}^{-\beta\mathcal{H}(t)}\delta(\mathsf{K}(t) - \mathsf{K}_0)}\right) - \overline{\Lambda}_t t \\ &= \ln\left(\mathrm{e}^{\beta[\mathcal{H}(t)-\mathcal{H}(0)]}\right) - \overline{\Lambda}_t t \\ &= \beta\int_0^t \dot{\mathcal{H}}(s)\ \mathrm{d}s - \overline{\Lambda}_t t.\end{aligned} \tag{3.26}$$

The form of the FT applicable in the isokinetic and canonical ensembles is therefore the same. For a Gauss isokinetic thermostatting mechanism, $\Lambda \equiv -d\mathrm{N}\alpha$, where α is the thermostat multiplier defined in Equation (2.43).

3.7 Lyapunov Instability

The stability of trajectories in phase space may be characterised by Lyapunov exponents. The stability of such trajectories is important with respect to the FT, as the propagation of trajectories through various phase volumes (§2.5.1) in phase space is affected by the Lyapunov instability of the trajectories.

[§]Note that the probability distribution function given here and in later sections is not normalised. Normalisation of the distribution function is not important in the derivation of the FT and is therefore ignored here to simplify the expressions in which it appears.

Figure 3.7: A depiction of the infinitesimal separation ($\delta\mathbf{\Gamma}$) of two trajectories in $\mathbf{\Gamma}$-space. After these trajectories are propagated for some time they noticeably diverge, which demonstrates that this model system is chaotic.

3.7.1 Lyapunov Exponents

Consider a trajectory in phase space that has a neighbouring trajectory that is separated from the main trajectory by an infinitesimal amount, as illustrated in Figure 3.7.

As the main trajectory is propagated through phase space, so too is the neighbouring trajectory. One of three things can happen at this point. The neighbouring trajectory will remain the same distance from the main trajectory (to within $\mathcal{O}(t^n)$), it will gradually diverge from the main trajectory, or it will gradually converge to the main trajectory. If the neighbouring trajectory diverges from the main trajectory, the system is said to be *chaotic*. The chaotic behaviour of the system may be expressed mathematically via the use of *Lyapunov exponents*. A system will have $2d$N Lyapunov exponents that characterise such behaviour.

If one defines $\mathbf{\Gamma}_0(0)$ as the initial point of the main phase space trajectory and $\mathbf{\Gamma}(0)$ as the starting point of a neighbouring trajectory, the separation vector, $\delta\mathbf{\Gamma}$, between these points is defined as $\delta\mathbf{\Gamma} = |\mathbf{\Gamma}(0) - \mathbf{\Gamma}_0(0)|$. It is then possible to characterise the rate at which the separation vector changes with time

$$\begin{aligned} \frac{\mathrm{d}\delta\mathbf{\Gamma}(t)}{\mathrm{d}t} &= \mathsf{T}(\mathbf{\Gamma}) \cdot \delta\mathbf{\Gamma}(t) \\ &= \frac{\partial\dot{\mathbf{\Gamma}}(\mathbf{\Gamma}(t))}{\partial\mathbf{\Gamma}} \cdot \delta\mathbf{\Gamma}(t) \end{aligned} \tag{3.27}$$

which implies that T is a matrix defined by

$$\mathsf{T} = \begin{pmatrix} \frac{\partial^2 \mathcal{H}}{\partial q \partial p} & \frac{\partial^2 \mathcal{H}}{\partial p^2} \\ -\frac{\partial^2 \mathcal{H}}{\partial q^2} & -\frac{\partial^2 \mathcal{H}}{\partial p \partial q} \end{pmatrix} \tag{3.28}$$

Using Equation (3.27), one can define a propagation matrix, $\mathsf{L}(t)$ of the form

$$\begin{aligned} \mathsf{L}(t) \cdot \delta\mathbf{\Gamma}(0) &= \mathrm{e}_{\mathrm{L}}^{\int_0^t \mathrm{d}s\, \mathsf{T}(\mathbf{\Gamma}(s))} \delta\mathbf{\Gamma}(0) \\ &= \delta\mathbf{\Gamma}(t) \end{aligned} \tag{3.29}$$

where e_{L} is the left time-ordered exponential. Equation (3.29) can be used to mathematically define the Lyapunov exponents of a system

$$\{\lambda_i; i = 1 \ldots 2d\mathrm{N}\} = \lim_{t\to\infty} \frac{1}{2t} \ln\Big[\text{eigenvalues}\big(\mathsf{L}(t)^{\mathsf{T}} \cdot \mathsf{L}(t)\big)\Big] \tag{3.30}$$

It is possible to use the theory associated with these exponents to formulate an alternate derivation of the Fluctuation Theorem. Such a derivation is given by Evans and Searles in Reference [44].

3.8 The Integrated FT

The integrated Fluctuation Theorem (IFT) was first proposed by Ayton, Evans and Searles in 2001 [45] during investigations into a local version of the FT. Mathematically, the IFT is written[¶]

$$\left[\frac{\Pr(\overline{\Omega}_t < 0)}{\Pr(\overline{\Omega}_t > 0)}\right] = \left\langle \mathrm{e}^{-\overline{\Omega}_t t} \right\rangle_{\overline{\Omega}_t > 0}. \tag{3.31}$$

That is, if one looks at the probability that the time-averaged entropy production will be positive to the probability it will be negative, this will be

[¶]Note that in this thesis we will use the notation $\frac{p_-}{p_+}$ interchangeably with $\frac{\Pr(\overline{\Omega}_t<0)}{\Pr(\overline{\Omega}_t>0)}$ when discussing the IFT.

related to the ensemble average of the exponential of the time-averaged entropy production only over those segments for which the time-averaged entropy production is positive. Equation (3.31) is essentially a statement of the Second Law, which is valid for finite systems for finite times.

In certain cases, verification of the Fluctuation Theorem, either in computer simulations or in an experimental laboratory, may be difficult due to insufficient statistics. For example, if one performs a computer simulation that is computationally expensive with the equipment one has available, it may not be feasible to allow the simulation to run for an extended amount of time. Likewise, it is possible that a laboratory experiment to test the FT could be extremely difficult or time consuming. In such cases, it is still possible to test the FT. However, instead of examining individual values for the time-averaged general dissipation function[‖] it is possible to look at the *sign* of $\overline{\Omega}_t$.

Armed with this new form of the FT, problems of insufficient statistics, as discussed above, can be overcome. The IFT also lends itself to investigations of causality, as writing computer programs that examine entropy response over a single positive and negative range is much simpler than attempting to examine entropy response over several different positive and negative values. More discussion in greater detail on this subject may be found in Chapter 8.

3.9 Generalised Fluctuation Theorem

Previously in this chapter the Fluctuation Theorem has been referred to as an expression that may be applied to time-averages of the dissipation function. It is, however, possible to obtain a FT for any given phase function. The derivation for a Generalised Fluctuation Theorem (GFT) [46] is almost identical to that given for the FT (§3.5). As there are some small differences, only an abbreviated derivation for the GFT follows.

[‖]see Equation (3.23).

3.9.1 Derivation of the GFT

We once again equate the probability ratios of finding a phase point in a given phase volume to the product of the ratio of distribution functions at a time $t = 0$ and exponential of the phase-space compression factor (Equation (3.20)). Again it is convenient to make use of a general dissipation function (§3.5.4) defined by Equation 3.22.

The exact form of Ω will depend on the ensemble in which one is interested. Such forms will be discussed in detail in chapters 4 through 8. The general form of the GFT for a phase function, $\mathcal{B}$, that is odd under time-reversal symmetry, is

$$\frac{\Pr(\overline{\mathcal{B}}_t = \mathrm{A})}{\Pr(\overline{\mathcal{B}}_t = -\mathrm{A})} = \left\langle \mathrm{e}^{-\overline{\Omega}_t t} \right\rangle^{-1}_{\overline{\mathcal{B}}_t = \mathrm{A}} \tag{3.32}$$

so the ratio of probabilities is equal to the inverse of the ensemble average of the exponentiated general dissipation function. Note that the ensemble average is only taken over those states for which $\overline{\mathcal{B}}_t = \mathrm{A}$.

3.10 FT and Green-Kubo relations

As mentioned briefly in the introduction to this chapter, the Fluctuation Theorem can be used to derive the Green-Kubo relations. Green-Kubo relations are important, as they can be used to calculate linear transport coefficients for a system [10]. The Green-Kubo relations were originally derived for non-equilibrium systems that were close to equilibrium. In their paper on the derivation of the Green-Kubo relations from the application of the Fluctuation Theorem [24], Searles and Evans noted that if a system is far from equilibrium, the phase-space distribution is not sufficiently Gaussian for the Green-Kubo relations to be satisfied. Thus the SSFT has been used to show that the Green-Kubo relations are valid only in the weak-field non-linear regime.

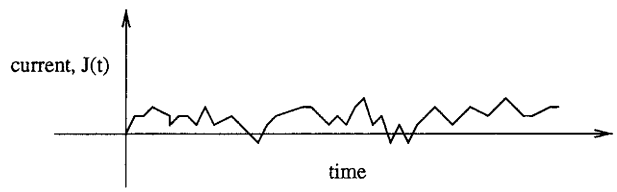


Figure 3.8: The response of the current of a *Gedanken* experiment over an arbitrary time interval. The Fluctuation Theorem predicts momentary (instantaneous) violations of the Second Law in microscopic systems

3.11 Numerical Tests

Many simulations have been done to test the validity of the FT [24, 41, 43, 45–54]. Simulations in the microcanonical, canonical and isokinetic ensembles have verified the applicability of the FT in all of these systems.

In order to numerically test the FT, it is not necessary to ever apply a time reversal mapping. By *only* running a simulation in the forward direction and monitoring the instantaneous, time-averaged and/or the ensemble-averaged values of a given phase function, it is possible to actually see these short-lived anti-trajectories occur spontaneously. Figure 3.8 shows a contrived example of what the instantaneous value of the current density, $J = \frac{\sum_i \mathbf{p}_i}{m} V^{-1}$, may look like in a *Gedanken* experiment where the current density is predicted to be positive by the Second Law of Thermodynamics and where the equations of motion are never mechanically reversed. In the figure, momentary violations of the Second Law are noted. While macroscopically this is forbidden (§3.2), these *microstates* are able to effect these momentary violations. In this *Gedanken* experiment, we see the behaviour that is predicted by the FT, including Second Law violations that are momentary and rare when compared to the average behaviour of the system.

3.11.1 Numerical Verification of the FT

As discussed earlier, the FT may be tested for any generic phase function. A prerequisite to this testing is, however, knowledge of how the phase function should normally respond to the presence of an external perturbing field. The irreversible entropy production, as previously mentioned, should respond with

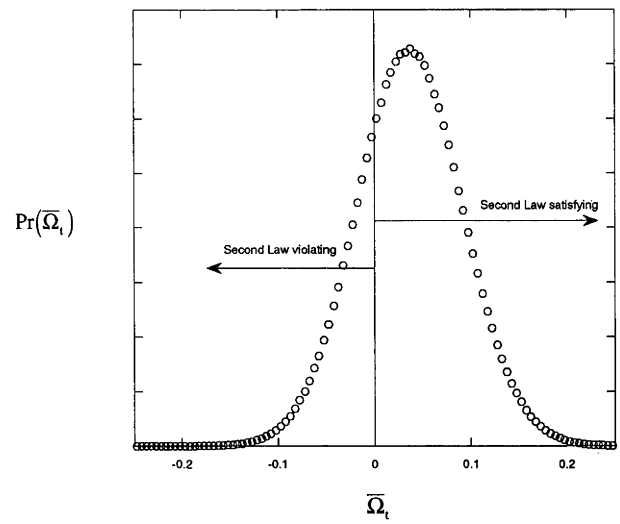


Figure 3.9: An illustration of the Second-law satisfying and Second-law violating parts of a probability histogram of a time-averaged phase function which is predicted to be positive by the Second Law of Thermodynamics.

a positive value in a Second Law satisfying system. Likewise, the current density of the system, $J = \frac{\sum_i \mathbf{p}_i}{m} V^{-1}$, should be positive if the system is Second Law satisfying.

By observing the time-averaged values of a phase function for each trajectory in a TFT MD simulation, it is possible to construct a probability histogram that gives a clear picture as to how frequently negative values for the current occur (i.e. how often the Second Law is violated). Figure 3.9 shows an illustration of one such possible histogram. Clearly, the Second Law violating states occur with much less frequency.

3.11.2 Nuances of MD FT Simulations

When actually writing molecular dynamics software that will test the Fluctuation Theorem, there are several items that must be considered to achieve the desired results.

Ensemble Averaging

One important consideration that is intrinsic to simulations testing the FT is ensemble averaging. In order to accurately record the correct information, it is necessary to store the sum of the values of the phase function in which one is

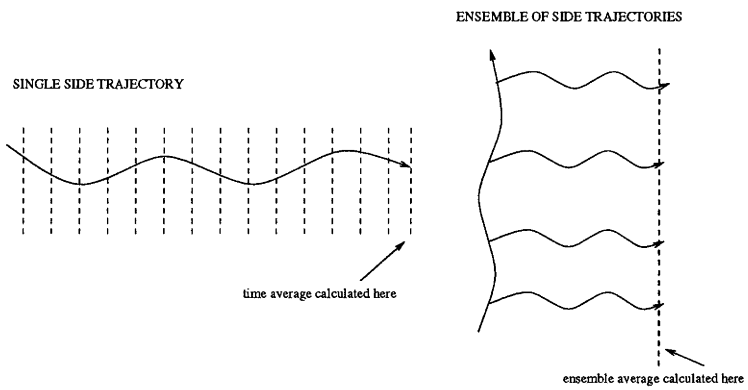


Figure 3.10: (left) The time average of a phase property is calculated by summing the values of the property at each step along the side trajectory. The sum is divided at the end by the number of steps, $\overline{\Psi}_t = \frac{\sum_{i=1}^{L} \Psi_i}{\sum_{i=1}^{L} 1}$. (right) The ensemble average is calculated by summing the time averages of multiple side trajectories, then dividing by the number of side trajectories, $< \overline{\Psi}_t >= \frac{\sum_{j=1}^{N} \overline{\Psi}_{tj}}{\sum_{j=1}^{N} 1}$.

interested at each step along the trajectory (either transient or steady-state). Figure 3.10 illustrates how one may calculate the time and ensemble averages in an MD simulation.

At the end of the trajectory, averaging of this sum over the trajectory length yields the time average of the phase function. The mathematical form of the time average is given in Equation (3.2). In an MD simulation, this integral is approached in the limit $\Delta t \rightarrow 0$, where Δt is the time step. In order to calculate the ensemble average, the sum of the time average of a phase function at the end of each trajectory (shown at the bottom of Figure 3.10 with a dashed line) is taken and divided by the number of trajectories over which the sum has been taken. The mathematical form of the ensemble average is given in Equation (3.3).

SSFT Simulations

When executing a simulation where the values of interest are in the steady-state of a transient trajectory, it is important to first determine when the system has reached a steady-state under the given conditions. This may be accomplished by performing some test simulations where the response of the phase function is monitored over the length of a transient trajectory segment and noting the

time at which the steady state is reached. Subsequent simulations with the same starting positions will yield a steady state at approximately the same time. One may then actually start the steady-state averaging of the phase function by storing only those values of the phase function after the trajectory has reached the steady state. Such averaging is done in a similar manner for each steady-state trajectory and ensemble averaging is then performed on each of these time averages as outlined above.

Equilibration Times

In order to ensure that the simulation takes place in the desired ensemble, it is necessary to allow an initial period at the beginning of the simulation where the system is allowed to equilibrate. This initial equilibration time of the backbone trajectory will vary depending on the conditions of the system, but testing can indicate what a reasonable value for such a time would be.

For the TFT and SSFT, another equilibration consideration is the amount of time that the system is allowed to evolve along the main trajectory before spawning a new transient trajectory. Normally this inter-trajectory equilibration time need not be nearly as long as the initial equilibration time at the beginning of the simulation. If t is too short there will be serial correlation between successive transient trajectories. t should be of the order of the time to relax to the steady state (i.e. a Maxwell time).

Chapter 4

Adiabatic Fluctuation Theorem

4.1 Introduction

As mentioned in Chapter 3, the Fluctuation Theorem has been considered in various ensembles. In this chapter, a new variation of the FT is examined, whereby the system starts from a canonical equilibrium distribution. The thermostat is then deactivated as an external field is imposed. The adiabatic response is then monitored for the time during which the system is subjected to the external field. This FT variant is denoted the adiabatic (i.e. no thermostat) Fluctuation Theorem.

The presence of an external field and the absence of a thermostat have two effects on the system. As in the general form of the FT discussed previously, the external field does work on the system and drives it away from equilibrium. By definition, because the system is adiabatic, the temperature of the system is no longer constrained. However, the phase-space compression factor, $\Lambda(\mathbf{\Gamma})$, is zero.

In all previous studies of the Fluctuation Theorem, the phase-space compression factor was non-zero. In (§3.5.3), the change in the phase-space volume elements over the duration of the propagation of a trajectory was discussed. Since the phase-space compression factor is zero in this version of the FT, the volume of the phase-space volume elements does not change over the course of the trajectory. This can be seen from the relationship (previously shown as

Equation (3.10))

$$\delta V(\mathbf{\Gamma}(t),t) = \delta V(\mathbf{\Gamma}(0),0)\mathrm{e}^{\overline{\Lambda}(t)t}. \tag{4.1}$$

The derivation of the adiabatic FT is presented in the next section. The methodology used in the MD simulations specifically designed to test this version of the FT is discussed in (§4.3). The results of these simulations are discussed in (§4.4).

4.2 Derivation of the Adiabatic FT

The derivation of this form of the Fluctuation Theorem is very similar to the general derivation shown in (§3.5). However, there are some differences between this system and those studied previously.

4.2.1 Adiabatic Hamiltonian

In (§2.2.1) we discussed the Hamiltonian and its relationship to the internal energy of a system at equilibrium where the two are equal. If, however, we have a system without a thermostat to which an external field is applied (Equation (2.7)), the change of the internal energy with time will be [10]

$$\begin{aligned}\dot{\mathcal{H}}_0^{ad}(\mathbf{\Gamma},t) &= \sum_{i=1}^{\mathrm{N}}\Big(\frac{\dot{\mathbf{p}}_i(t)\cdot\mathbf{p}_i}{m} - \dot{\mathbf{q}}_i(t)\cdot\mathbf{F}_i\Big)\\ &= -\sum_{i=1}^{\mathrm{N}}\Big(-\mathbf{D}_i\cdot\frac{\mathbf{p}_i}{m} + \mathbf{C}_i\cdot\mathbf{F}_i\Big)\mathrm{F}_e(t)\\ &\equiv -\mathrm{J}(\mathbf{\Gamma})\mathrm{V}\mathrm{F}_e(t).\end{aligned} \tag{4.2}$$

Here we have defined the *dissipative* flux as

$$\mathrm{J}(\mathbf{\Gamma})\mathrm{V} \equiv \sum_{i=1}^{\mathrm{N}}\Big(-\mathbf{D}_i\cdot\frac{\mathbf{p}_i}{m} + \mathbf{C}_i\cdot\mathbf{F}_i\Big). \tag{4.3}$$

$\mathbf{C}_i$ and $\mathbf{D}_i$ represent the coupling of the external field to the system.

4.2.2 Adiabatic FT Derivation

It is now possible to continue with the adiabatic Fluctuation Theorem derivation. From (§2.6.2), we know that the phase-space distribution in the canonical ensemble at time $t = 0$ of any given phase function at any time t is

$$f(\mathbf{\Gamma}(t), 0) = \mathrm{e}^{-\beta_0 \mathcal{H}(\mathbf{\Gamma}(t),0)}. \tag{4.4}$$

Using the definition of the general dissipation function (Equation (3.22)) with no phase-space compression, yields

$$\overline{\Omega}_t t \equiv \int_0^t \mathrm{d}s\, \Omega(\mathbf{\Gamma}(s)) = \ln\Big(\frac{f(\mathbf{\Gamma}(0), 0)}{f(\mathbf{\Gamma}(t), 0)}\Big). \tag{4.5}$$

We then have

$$\begin{aligned} \overline{\Omega}_t t &= \ln\Big[\frac{f(\mathbf{\Gamma}(0), 0)}{f(\mathbf{\Gamma}(t), 0)}\Big] \\ &= \ln\Big[\frac{\mathrm{e}^{-\beta_0 \mathcal{H}_0(\mathbf{\Gamma}(0))}}{\mathrm{e}^{-\beta_0 \mathcal{H}_0(\mathbf{\Gamma}(t))}}\Big] \\ &= \ln\big(\mathrm{e}^{\beta_0[\mathcal{H}_0(\mathbf{\Gamma}(t)) - \mathcal{H}_0(\mathbf{\Gamma}(0))]}\big) \\ &= \beta_0 \int_0^t \mathrm{d}s\, \dot{\mathcal{H}}_0^{ad}(\mathbf{\Gamma}(s)) \end{aligned} \tag{4.6}$$

Combining Equations (4.2) and (4.6) yields the relation

$$\begin{aligned} \overline{\Omega}_t &= -\beta_0 \frac{1}{t} \int_0^t \mathrm{d}s\, \mathrm{J}(s)\mathrm{VF}_e \\ &= -\beta_0 \overline{\mathrm{J}}_t \mathrm{VF}_e. \end{aligned} \tag{4.7}$$

Here $\beta_0 = \frac{1}{\mathrm{k_B T_0}}$ and $\mathrm{T_0}$ is the absolute temperature at $t = 0$, not the time-evolved temperature.

We can now substitute this new expression for the general dissipation func-

tion into the master FT equation (Equation (3.23)). Doing so yields

$$\ln\left(\frac{\Pr(\overline{\Omega}_t = \mathrm{A})}{\Pr(\overline{\Omega}_t = -\mathrm{A})}\right) = \ln\left(\frac{\Pr(-\beta_0 \overline{\mathrm{J}}_t \mathrm{VF}_e = \mathrm{A})}{\Pr(-\beta_0 \overline{\mathrm{J}}_t \mathrm{VF}_e = -\mathrm{A})}\right) = \mathrm{A}t \tag{4.8}$$

A non-thermostatted system therefore obeys the FT given by Equation (4.8). [50] In order to verify this, it is possible to perform MD simulations that test this equation. From the Second Law of Thermodynamics, it is expected that as an external field is applied, the system will respond in a particular direction such that $\langle \dot{\mathcal{H}}_0^{ad} \rangle > 0$. Equation (4.8) is consistent with the Second Law and allows one to determine the probability of observing Second-law violating states (e.g. dissipative flux flowing in the direction opposite to what is expected) for a finite system in a finite time in the absence of a thermostat.

For purposes of discussing the MD simulations and results in the sections that follow, we will use the terminology *remnant entropy production* for $\overline{\Omega}_t$, as defined in Equation (4.7). The expression for the irreversible entropy production is

$$\begin{aligned} \frac{\overline{\Sigma}_t}{\mathrm{k_B}} &= -\left[\overline{\frac{\mathrm{J}(t)}{\mathrm{T}(t)}}\right]_t \frac{\mathrm{VF}_e}{\mathrm{k_B}} \\ &= -[\overline{\beta(t)\mathrm{J}(t)}]_t \mathrm{VF}_e. \end{aligned} \tag{4.9}$$

Note that this differs from the expression for the remnant entropy production by the temperature that is used. For the irreversible entropy production, it is the instantaneous temperature, $\mathrm{T}(t)$, while for the remnant entropy production, it is the temperature at time $t = 0$, T_0.

4.3 Adiabatic MD Simulations

Several MD simulations were performed to test Equation (4.8). Since we want our dynamics to be canonical-ensemble generating, in order to test the derivation outlined in the previous section by creating an initial canonical distribution, we make use of the Nosé-Hoover thermostatting mechanism (§2.9.3).

Our plan is to equilibrate a system that is described by a canonical distri-

bution function. Once this equilibrium state is reached, we can then spawn transient trajectories from this equilibrium trajectory. These transient trajectories differ from the equilibrium trajectory in two ways. First, they are influenced by the external field and second, they have no thermostat. The non-equilibrium system we study undergoes adiabatic colour diffusion (§2.8.1).

To accomplish this, we need two different sets of the equations of motion. For the equilibrium trajectory, the equations of motion are

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \zeta \mathbf{p}_i, \end{aligned} \tag{4.10}$$

where ζ is the Nosé-Hoover thermostat multiplier. The non-equilibrium equations of motion are

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i + \mathbf{i}c_i \mathrm{F}_e, \end{aligned} \tag{4.11}$$

where $c_i = (-1)^i$ is the colour coupling constant and F_e is the external colour field. For this system, the dissipative flux is $\mathrm{J}_x = -\frac{1}{\mathrm{V}} \sum_i^{\mathrm{N}} c_i \mathbf{p}_{xi}$ and the current density is defined in Equation (2.29).

All simulations performed for these investigations had the following conditions: N = 32, n = 0.4, time step=0.004, T_0 = 1.0, WCA potential. The mass of the heat bath used for the NH thermostatting mechanism was Q = 10. The equilibration time was 10,000 time steps and the transient equilibration time (i.e. the time allowed for equilibration between propagation of successive adiabatic segments) was 200 time steps. The external field strength and length of the transient trajectories varied, depending on the simulation.

4.4 Results

The field strengths were varied in order to examine the influence of the external field on the test system and its adherence to the adiabatic FT expression. The

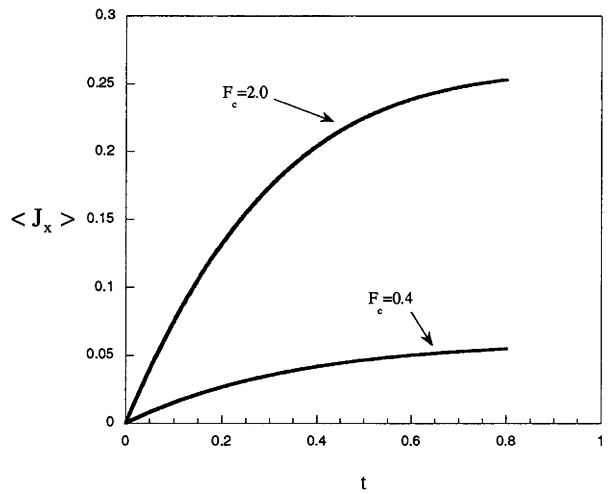


Figure 4.1: The ensemble-averaged current density for an MD simulation that tests the adiabatic FT. Two separate simulations were performed with two different field strengths ($F_c = 0.4, 2.0$). Note that the response to the stronger field is much greater, due to the increased amount of work done on the system by the field.

transient trajectory lengths were varied to illustrate the dependence of the probability distribution of the dissipative flux on the averaging times.

Figure 4.1 shows the response of the ensemble-averaged current density with time over a transient trajectory of length $t_{max} = 0.8$ for two different field strengths. The external fields used were $F_c = 0.4$ and $F_c = 2.0$. The ensemble average of the current density is much greater for the higher field strength, due to the greater amount of work done by the external field on the system.

Figure 4.2 shows the change in the temperature of the system over the length of a transient trajectory for two different field strengths ($F_c = 0.4, 2.0$) used in two different MD simulations. The temperature increases much faster when the system is exposed to a stronger external field, as more work is done on the system, and for small fields the ensemble average of the work done is proportional to F_c^2. It is possible that for the higher (i.e. $F_c = 2.0$) field strength the rate at which work is done on the system is proportional to terms of even higher order. This causes the system to heat-up at a faster rate.

Figure 4.3 shows a comparison of MD simulation data that test the master adiabatic FT equation (Equation (4.8)) with $F_c = 0.4$. Two sets of data are shown. The black crosses show a test of this equation using the remnant

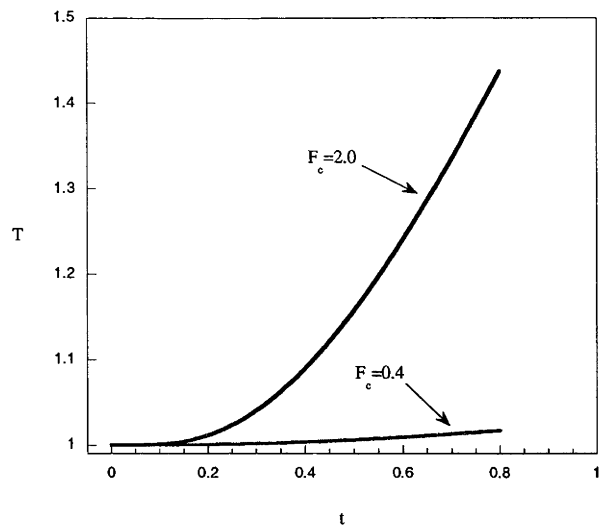


Figure 4.2: The change in temperature over the length of a transient trajectory for two separate MD simulations with different external field strengths (F_c = 0.4, 2.0). The temperature increases much more rapidly for the system with a stronger external field. This is caused by the system heating-up at a faster rate due to more work being done on the system.

entropy production, $\overline{\Omega}_t$, while the red circles are the results of the same test using the irreversible entropy production, $\frac{\overline{\Sigma}_t}{k_B}$. As expected, the test using the remnant entropy production satisfies and verifies the adiabatic Fluctuation Theorem, while the irreversible entropy production clearly does not satisfy this fluctuation relation. In order to have a better understanding of exactly why this occurs, we can look again at Figure 4.2. Notice how the temperature changes over the course of the propagation of the transient trajectory. Clearly, the use of $T(t)$, rather than T_0, will yield results that are not in agreement with what is predicted by the Fluctuation Theorem.

Figure 4.4 illustrates the effect of the transient trajectory length on the probability distribution of the dissipative flux. The distribution for four different transient trajectory lengths is shown from simulations with $F_c = 0.4$. As the length of the transient trajectory increases, the standard deviation of the probability distribution decreases due to the longer averaging time. Additionally, the average value of the remnant entropy production increases due to the work done by the field. Longer simulation times are therefore required to test the FT with longer transient trajectory lengths, as fewer Second-law violating states are observed. This can be seen in Figure 4.4.

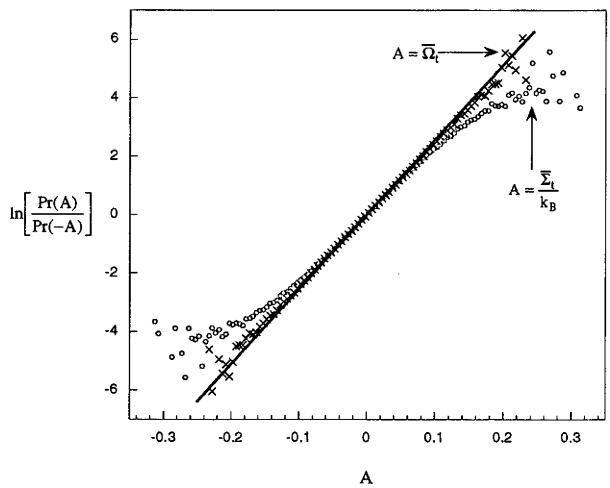


Figure 4.3: A comparison of the results of testing the adiabatic FT equation (Equation (4.8)) for the irreversible entropy production ($\frac{\overline{\Sigma}_t}{k_B}$, red circles) and the remnant entropy production ($\overline{\Omega}_t$, black crosses) with $F_c = 0.4$. The blue line is what the adiabatic FT predicts. Note that the expression testing the FT using the remnant entropy production clearly adheres to what is predicted by the adiabatic FT. The test of the FT using the irreversible entropy production, however, does not work, as is expected.

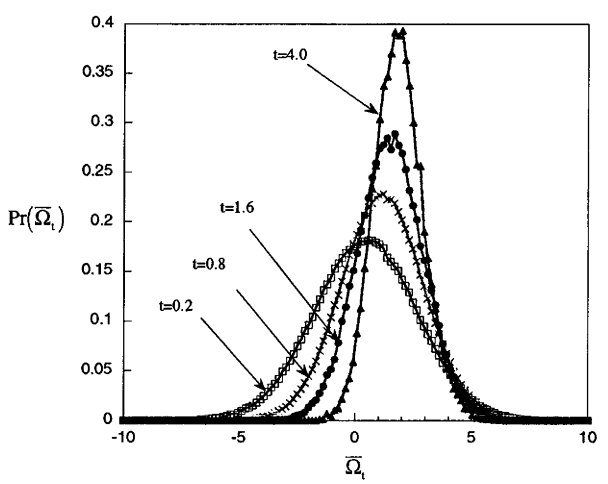


Figure 4.4: The probability histograms for four MD simulations with different transient trajectory lengths with $F_c = 0.4$. As the length of the transient trajectory increases, it can be seen that the standard deviation of the histogram decreases. This is due to the longer averaging time over the length of the transient trajectory. The average value of the remnant entropy production is seen to increase as the transient trajectory length increases. This is due to the work done by the field.

4.5 Summary

We have shown that the Fluctuation Theorem is valid for systems in the canonical ensemble with the absence of a thermostat. We have also examined the effects of varying the lengths of the trajectories over which the dissipative flux is averaged. Finally, we have shown that the irreversible entropy production cannot be used as the dissipation function, Ω, in the Fluctuation Theorem for this system.

Chapter 5

FT for Hamiltonian Relaxation

5.1 Introduction

In Chapters 3 and 4, several forms of the Fluctuation Theorem have been discussed. In this chapter, we will examine yet another application of the FT. In this variation, instead of beginning with a system at equilibrium, as in all previous versions of the FT, we begin with a system subject to an external field. Periodically, the field is turned off. The moment the non-dissipative field is turned off, the system is no longer in a non-equilibrium steady state due to the change in the Hamiltonian. The system is then monitored as it relaxes towards equilibrium [50].

In this chapter, we consider a system that is initially thermostatted and under the influence of a time independent sinusoidal external colour field. This system is allowed to reach a steady state. From this main trajectory, side trajectories are periodically spawned. When a new side trajectory is spawned, the thermostat and external field are switched off, allowing the system to relax under entirely Newtonian dynamics and the coloured particles then mix with one another. At the end of the side trajectory, the coordinates from the beginning of the side trajectory are read back in. The thermostat and external field are reactivated and the main trajectory is once again propagated for a given number of time steps until another side trajectory is spawned.

5.2 Hamiltonian Relaxation FT

5.2.1 System Hamiltonian

For the Hamiltonian relaxation FT, the initial Hamiltonian will consist of the interaction Hamiltonian, $\mathcal{H}_0$, and a term for the external field. The initial Hamiltonian is

$$\mathcal{H} = \mathcal{H}_0 + \sum_i c_i \mathrm{F}_c \sin(k\mathrm{x}_i), \tag{5.1}$$

where F_c is the external colour field, c_i is the colour coupling constant and x_i is the position of particle i in the x direction. At time $t < 0$, $\mathcal{H}$ is a constant of motion, while at time $t \geq 0$, $\mathcal{H}_0$ is constant. The change of the system Hamiltonian with time is then

$$\begin{aligned} \frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} &= \frac{\mathrm{d}\mathcal{H}_0}{\mathrm{d}t} + \sum_i^{\mathrm{N}} c_i \mathrm{F}_c k\dot{x} \cos(kx_i) \\ &= \sum_i^{\mathrm{N}} c_i \mathrm{F}_c k\dot{x} \cos(kx_i), \qquad t \geq 0. \end{aligned} \tag{5.2}$$

5.2.2 Initial Ensemble

For time $t < 0$, the phase-space distribution of the system is

$$f(\mathbf{\Gamma}(0), 0) = \mathrm{e}^{-\beta(\mathrm{H}_0 + \sum_i c_i \mathrm{F}_c \sin(k\mathrm{x}_i))} \tag{5.3}$$

For $t > 0$, the external field is deactivated so we may observe the diffusion of the two groups of ("coloured") particles. The equations of motion are:

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \mathrm{F}_c(t) k c_i \cos(k\mathrm{x}_i)\mathbf{i}_x - \zeta \mathbf{p}_i (1 - \Theta(t)) \\ \mathrm{F}_c(t) &= \mathrm{F}_c(1 - \Theta(t)), \end{aligned} \tag{5.4}$$

where Θ is the Heaviside step function.

The general dissipation function for this version of the FT is then

$$\overline{\Omega}_t t = \ln\left[\frac{f(\mathbf{\Gamma}(0),0)}{f(\mathbf{\Gamma}(t),0)}\right]. \tag{5.5}$$

The phase-space compression factor $\Lambda \equiv 0$, since there is no thermostat present in the side trajectories. Note that the presence of a thermostat in the main trajectory is only important in ensuring that the initial ensemble is canonical.

We then have

$$\begin{aligned}\overline{\Omega}_t t &= \beta\left[\mathcal{H}(\mathbf{\Gamma}(t)) - \mathcal{H}(\mathbf{\Gamma}(0))\right] \\ &= \beta \int_0^t \mathrm{d}s\, \dot{\mathcal{H}}(s).\end{aligned} \tag{5.6}$$

Substituting Equation (5.2) into Equation (5.6) yields

$$\overline{\Omega}_t t = \beta \mathrm{F}_c k \int_0^t \mathrm{d}s \sum_i^{\mathrm{N}} c_i \dot{x}_i(s) \cos(k x_i(s)). \tag{5.7}$$

We can then use this definition of the general dissipation function in the master FT equation to obtain the colour diffusion FT expression

$$\ln\left(\frac{\Pr(\overline{\Omega}_t = \mathrm{A})}{\Pr(\overline{\Omega}_t = -\mathrm{A})}\right) = \ln\left(\frac{\Pr(\beta \mathrm{F}_c k \sum_i^{\mathrm{N}} c_i \overline{\left[p_{xi}\cos(kx_i)\right]}_t = \mathrm{A})}{\Pr(\beta \mathrm{F}_c k \sum_i^{\mathrm{N}} c_i \overline{\left[p_{xi}\cos(kx_i)\right]}_t = -\mathrm{A})}\right) = \mathrm{A}t. \tag{5.8}$$

5.3 MD Simulations

In the simulations we performed, we used a small number of particles ($\mathrm{N} = 8$). We chose to do this for two reasons. First, for large number of particles, it becomes difficult or impossible to observe the Second-law violating fluctuations. Second, we wanted to have a sufficient number of particles to be able to verify the FT, yet not so many as to make the computational expense of the calculations prohibitive.

Since we used a small number of particles, it is important to consider any $\mathcal{O}(\frac{1}{\mathrm{N}})$ effects (§2.7.2). The conditions imposed on the system by the restrictions outlined in the introduction of this chapter will influence the number of degrees

of freedom. For the main trajectory, the number of degrees of freedom will be reduced by one due to constraints in the momenta in the y direction. In the two-dimensional system used in our MD simulations, the number of degrees of freedom in momentum space is $2N - 1$. The momenta in the x direction are not conserved and the number of degrees of freedom are not further reduced in this direction. The thermostat does not reduce the number of degrees of freedom for the main trajectory. The Nosé-Hoover thermostatting mechanism is used, which does not constrain the instantaneous value of the kinetic energy, as previously discussed in (§2.9.3).

For the side trajectories, the number of degrees of freedom differs from that of the main trajectory. The absence of an external field on the side trajectories results in the momenta in both the x and y directions being conserved. The absence of a thermostat on the side trajectories has no influence on the number of degrees of freedom. The number of degrees of freedom for the side trajectories is therefore $2N - 2$.

For the results of the MD simulation discussed later in this chapter, the conditions of the system were $N = 8$, $n = 0.3$, $k = \frac{2\pi}{L}$, $F_c = 0.15$, time step=0.004 and WCA potential. Periodic boundary conditions were used. The main trajectory utilised a Nosé-Hoover thermostat with a mass of the heat bath of $Q = 10$. The ensemble averages were computed over 1,056,000 side trajectories.

5.4 Results

A single MD simulation was performed to test the FT for Hamiltonian relaxation. Figure 5.1 shows a probability histogram of the dissipation function for this system. As expected, the majority of the averages are positive (Second Law satisfying).

Figure 5.2 shows the result of a plot of the MD simulation data that directly tests Equation (5.8). The plotted data are depicted as black circles, while the red line is that predicted by the FT for Hamiltonian relaxation. Excellent agreement is observed, indicating that this version of the FT has been

confirmed.

5.5 Summary

We have presented another application of the Fluctuation Theorem that involves a system subjected to an external field that is later allowed to relax while the dissipative flux is monitored. The results of molecular dynamics simulations indicate that the FT is verified for this case, as expected.

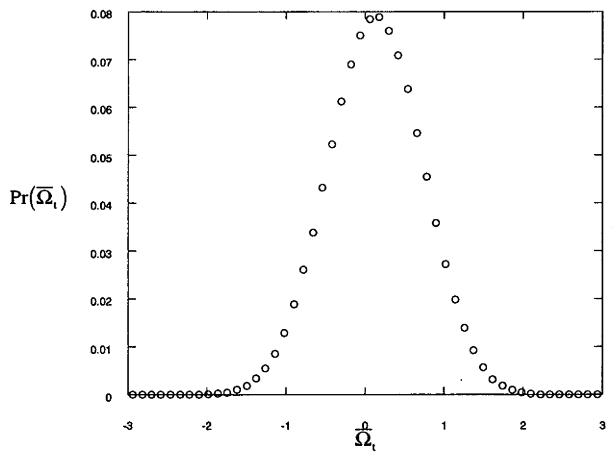


Figure 5.1: The probability histogram of the dissipative flux for an MD simulation. The conditions of the simulated system were N = 8, $n = 0.3$ $k = \frac{2\pi}{L}$, $F_c = 0.15$, time step=0.004 and WCA potential. Periodic boundary conditions were used. The main trajectory is thermostatted using a Nosé-Hoover mechanism. The data were obtained from a simulation with 1,056,000 side trajectories. Here the subscript t is the length of the side trajectories.

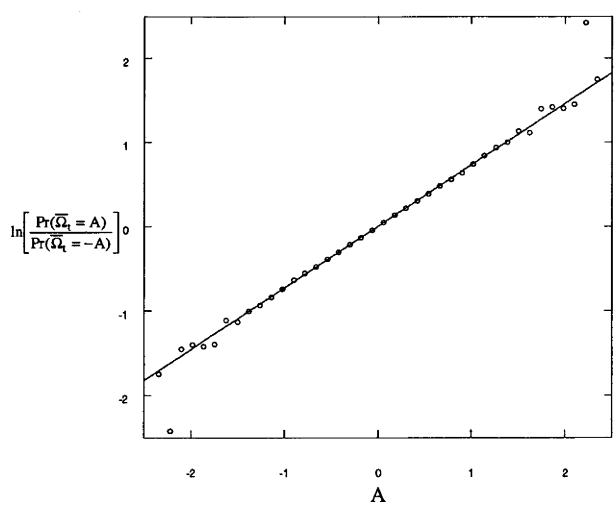


Figure 5.2: A test of the colour diffusion FT. The results of the MD simulation are shown as black circles, while the line predicted by the FT for Hamiltonian relaxation is shown in red. This application of the FT is shown to be verified, as expected. The data plotted here are the same as those used for the probability histogram in Figure 5.1.

Chapter 6

Isothermal-isobaric Formulation of the Fluctuation Theorem

6.1 Introduction

Previously, discussions of the Fluctuation Theorem have focused primarily on isochoric systems [43]. In this chapter, we present an application of the Fluctuation Theorem to systems in the isothermal-isobaric ensemble [55]. This ensemble presents many interesting problems that must be considered in order to verify the FT at constant pressure and temperature. These problems are not only theoretical ones, as dealt with in the derivation, but also practical ones with implementing isothermal-isobaric MD simulations.

In this chapter we first discuss the isothermal-isobaric ensemble. The derivation of the Fluctuation Theorem in this ensemble then follows. We then will discuss the nuances of MD simulations in this ensemble that verify the FT. We examine a particle interaction potential not previously considered in this thesis. We will also examine the suitability of the WCA and Lennard-Jones potentials for constant pressure, constant temperature dynamics. Finally, we present results of MD simulations and demonstrate that the FT is indeed applicable in this ensemble.

6.2 Isothermal-isobaric Ensemble

In this work we consider systems that have a constant kinetic energy and constant hydrostatic pressure. As in the isokinetic ensemble, a thermostat is used to hold the kinetic temperature constant. In order to hold the hydrostatic pressure constant, a barostat mechanism, similar to the thermostat mechanism, must be implemented. This introduces new problems into the design of MD simulations, which are discussed below.

The phase-space distribution of the initial equilibrium ensemble is

$$f(\Gamma, V) \sim \delta(p - p_0)\delta(K - K_0)e^{-\beta_0(\mathcal{H}_0 + p_0 V)}, \tag{6.1}$$

where p_0, K_0 are the fixed values of the pressure and kinetic energy, β_0 is the Boltzmann factor for the initial temperature of the system, V is the volume of the N-particle system and p is the hydrostatic pressure given by

$$\begin{aligned} pV &= \frac{1}{d}\Big(\sum_i \mathbf{p}_i \cdot \mathbf{p}_i + \sum_i \mathbf{q}_i \cdot \mathbf{F}_i\Big) \\ &= \frac{1}{d}\Big(\sum_i \mathbf{p}_i \cdot \mathbf{p}_i - \frac{1}{2}\sum_{ij} \mathbf{q}_{ij} \cdot \mathbf{F}_{ij}\Big), \end{aligned} \tag{6.2}$$

where $\mathbf{q}_{ij} = \mathbf{q}_j - \mathbf{q}_i$.

6.2.1 Molecular Dynamics Implementation

To maintain the temperature at a constant value, we use a Gaussian thermostatting mechanism. Constant hydrostatic pressure is maintained with a Gaussian barostat mechanism that functions very similarly to the Gaussian thermostat mechanism. By maintaining the pressure at a constant value, the volume must be allowed to fluctuate. The Gaussian barostat makes the pressure a constant of the motion, i.e. the *instantaneous* pressure is held constant (to within a numerical error threshold in computer simulations).

Results that are very similar would be obtained via the use of a Nosé-Hoover barostat mechanism. However, this method allows fluctuations in the pressure

and the volume at the same time (i.e. the *average* pressure is constant). Previous studies have shown that the Gaussian barostat mechanism, while more difficult to implement in a computer program than the corresponding Nosé-Hoover barostat mechanism, is generally more stable [10].

In order to force the system out of equilibrium, colour diffusion is used (§2.8.1).

6.2.2 Equations of Motion

The equations of motion for the equilibrium isothermal-isobaric system are

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} + \dot{\varepsilon}\mathbf{q}_i \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \dot{\varepsilon}\mathbf{p}_i - \alpha\mathbf{p}_i \\ \dot{\mathrm{V}} &= d\mathrm{V}\dot{\varepsilon}. \end{aligned} \tag{6.3}$$

Here $\dot{\varepsilon}$ is the *dilation rate*, which describes how the system volume expands or contracts as the pressure is maintained at a constant value. This is most easily illustrated if we consider a system where $\mathbf{p}_i = 0\ \forall i$ and the particles do not interact. Then the equations of motion become

$$\dot{\mathbf{q}}_i = \dot{\varepsilon}\mathbf{q}_i. \tag{6.4}$$

The mathematical expression for the dilation rate in Equation (6.3), that keeps the pressure constant is

$$\dot{\varepsilon} = -\frac{\frac{1}{2m}\sum_{i\neq j}\mathbf{q}_{ij}\cdot\mathbf{p}_{ij}\left(\phi''_{ij} + \frac{\phi'_{ij}}{\mathrm{q}_{ij}}\right)}{\frac{1}{2}\sum_{i\neq j}\mathrm{q}_{ij}^2\left(\phi''_{ij} + \frac{\phi'_{ij}}{\mathrm{q}_{ij}} + d^2\mathrm{pV}\right)} \tag{6.5}$$

and the form of the Gaussian thermostat multiplier is

$$\alpha = -\dot{\varepsilon} + \frac{\sum_{i=1}^{\mathrm{N}}\mathbf{F}_i\cdot\mathbf{p}_i}{\sum_{i=1}^{\mathrm{N}}\mathbf{p}_i^2}. \tag{6.6}$$

6.2.3 System Hamiltonian

If we consider our system to be contained in a d-dimensional cube with a side of length L, the volume of our system is then L^d. As the pressure is maintained at a constant value, the volume changes according to the third equation of motion in Equation (6.3),

$$\dot{\mathrm{V}} = d\mathrm{V}\dot{\varepsilon}. \tag{6.7}$$

The internal energy of the system is

$$\mathcal{H}_0 = \sum_{i=1}^{\mathrm{N}} \frac{\mathbf{p}_i^2}{m} + \Phi(\mathbf{q}). \tag{6.8}$$

The change in the internal energy with time is

$$\begin{aligned} \dot{\mathcal{H}}_0 &= \sum_{i=1}^{\mathrm{N}} \left(\frac{\mathbf{p}_i}{m} \cdot \dot{\mathbf{p}}_i - \mathbf{F}_i \cdot \dot{\mathbf{q}}_i\right) \\ &= \sum_{i=1}^{\mathrm{N}} -\dot{\varepsilon}\frac{\mathbf{p}_i^2}{m} + \dot{\varepsilon}\mathbf{q}_i\mathbf{F}_i - \alpha\frac{\mathbf{p}_i^2}{m} \\ &= -d\dot{\varepsilon}\mathsf{p}\mathrm{V} - 2\mathsf{K}\alpha, \end{aligned} \tag{6.9}$$

since $\frac{\mathrm{p}^2}{m} = 2\mathsf{K}$.

From the Equipartition Theorem, $2\mathsf{K} = d\mathrm{N}\mathrm{k}_\mathrm{B}\mathrm{T}$. Using Equation (6.7), the time derivative of the system Hamiltonian can be rewritten as

$$\dot{\mathcal{H}}_0 = -\mathsf{p}_0\dot{\mathrm{V}} - d\mathrm{N}\mathrm{k}_\mathrm{B}\mathrm{T}\alpha. \tag{6.10}$$

6.2.4 Phase-space Compression

The phase-space compression factor, $\Lambda(\mathbf{\Gamma})$ (§2.5.1) for the isothermal-isobaric ensemble is

$$\begin{aligned}\Lambda = \frac{\partial(\dot{\Gamma}, \dot{\mathrm{V}})}{\partial(\Gamma, \mathrm{V})} &= \sum_{i=1}^{\mathrm{N}} \frac{\partial \dot{\mathbf{q}}_i}{\partial(\Gamma(\mathbf{p}_i, \mathbf{q}_i), \mathrm{V})} + \sum_{i=1}^{\mathrm{N}} \frac{\partial \dot{\mathbf{p}}_i}{\partial(\Gamma(\mathbf{p}_i, \mathbf{q}_i), \mathrm{V})} + \frac{\partial \dot{\mathrm{V}}}{\partial \Gamma, \mathrm{V}} \\ &= d\mathrm{N}\dot{\varepsilon} - d\mathrm{N}\dot{\varepsilon} - d\mathrm{N}\alpha + \mathcal{O}(1) \\ &= -d\mathrm{N}\alpha + \mathcal{O}(1)\end{aligned} \tag{6.11}$$

In order to tell if the equilibrium phase-space distribution (Equation (6.1)) will change with the application of the isothermal-isobaric dynamics, we evaluate the Liouville equation (§2.21). However, to evaluate this expression, we use Equation (6.10), which can be written

$$\frac{\mathrm{d}(\mathcal{H}_0 + \mathsf{p}_0 \mathrm{V})}{\mathrm{d}t} = -d\mathrm{N}\mathrm{k}_\mathrm{B}\mathrm{T}_0\alpha. \tag{6.12}$$

The Liouville equation is then

$$\begin{aligned}\frac{\partial f_{eq}(\Gamma, \mathrm{V})}{\partial t} &= -f_{eq}\frac{\partial}{\partial(\Gamma, \mathrm{V})} \cdot (\dot{\Gamma}, \dot{\mathrm{V}}) - (\dot{\Gamma}, \dot{\mathrm{V}}) \cdot \frac{\partial f_{eq}}{\partial(\Gamma, \mathrm{V})} \\ &= \beta_0(\dot{\mathcal{H}}_0 + \mathsf{p}_0\dot{\mathrm{V}})f_{eq} + d\mathrm{N}\alpha f_{eq} \\ &= -d\mathrm{N}\beta_0\mathrm{k}_\mathrm{B}\mathrm{T}_0\alpha f_{eq} + d\mathrm{N}\alpha f_{eq} \\ &= 0,\end{aligned} \tag{6.13}$$

so the equilibrium phase-space distribution is indeed preserved under this dynamics.

6.2.5 Non-equilibrium Isothermal-isobaric Systems

We now look at the effects of introducing an external field on the equations of motion. We now incorporate terms into the equations of motion for the

coupling of the field to the generalised coordinates and momenta

$$\begin{aligned}\dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} + \dot{\varepsilon}\mathbf{q}_i + \mathbf{C}_i\mathrm{F}_c \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \dot{\varepsilon}\mathbf{p}_i + \mathbf{D}_i\mathrm{F}_c - \alpha\mathbf{p}_i \\ \dot{\mathrm{V}} &= d\mathrm{V}\dot{\varepsilon}\end{aligned} \tag{6.14}$$

For simplicity, we use colour diffusion (§2.8.1). Other algorithms for simulating external fields could also be used with equal results (e.g. SLLOD). The equations of motion with an external colour field are

$$\begin{aligned}\dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} + \dot{\varepsilon}\mathbf{q}_i \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \dot{\varepsilon}\mathbf{p}_i + \mathbf{i}c_i\mathrm{F}_c - \alpha\mathbf{p}_i \\ \dot{\mathrm{V}} &= d\mathrm{V}\dot{\varepsilon}.\end{aligned} \tag{6.15}$$

The form of the dilation rate remains the same as in equilibrium case and the Gaussian thermostat multiplier in the isothermal-isobaric ensemble with an external colour field present is

$$\alpha = -\dot{\varepsilon} + \frac{\sum_{i=1}^{\mathrm{N}}(\mathbf{F}_i - \mathbf{i}c_i\mathrm{F}_c)\cdot\mathbf{p}_i}{\sum_{i=1}^{\mathrm{N}}\mathbf{p}_i\cdot\mathbf{p}_i}. \tag{6.16}$$

The enthalpy of a system in the isothermal-isobaric ensemble is defined by [10]

$$\mathrm{I} \equiv \mathcal{H}_0 + \mathrm{p}_0\mathrm{V}. \tag{6.17}$$

The dissipative flux (§4.3) of such a system is then given by

$$\dot{\mathrm{I}}^{ad} \equiv -\mathbf{J}\mathrm{V}\cdot\mathbf{F}_c. \tag{6.18}$$

6.3 Derivation of the Isothermal-isobaric FT

We now have the necessary tools to complete a derivation of the isothermal-isobaric FT. Since the phase-space compression of this ensemble is non-zero,

this term must be incorporated into the expression for the general dissipation function, unlike the versions of the FT examined in the previous two chapters.

The general dissipation function is then

$$\begin{aligned}
\int_0^t \mathrm{d}s\ \Omega(\mathbf{\Gamma}(s)) &= \ln\Big[\frac{f_{eq}(\mathbf{\Gamma}(0),\mathrm{V}(0),0)}{f_{eq}(\mathbf{\Gamma}(t),\mathrm{V}(t),0)}\Big] - \int_0^t \mathrm{d}s\ \Lambda(\mathbf{\Gamma}(s)) \\
&= \ln\Big[\frac{\delta(\mathsf{p}(0)-\mathsf{p}_0)\delta(\mathsf{K}(0)-\mathsf{K}_0)\mathrm{e}^{-\beta_0(\mathcal{H}_0+\mathsf{p}_0\mathrm{V})}}{\delta(\mathsf{p}(t)-\mathsf{p}_0)\delta(\mathsf{K}(t)-\mathsf{K}_0)\mathrm{e}^{-\beta_0(\mathcal{H}_0+\mathsf{p}_0\mathrm{V})}}\Big] + \int_0^t \mathrm{d}s\ d\mathrm{N}\alpha(s) \\
&= \beta_0[(\mathcal{H}_0(t)+\mathsf{p}_0\mathrm{V}(t)) - (\mathcal{H}_0(0)+\mathsf{p}_0\mathrm{V}(0))] + \int_0^t \mathrm{d}s\ d\mathrm{N}\alpha(s) \\
&= -\int_0^t \mathrm{d}s\ \beta_0\mathbf{J}(s)\mathrm{V}(s)\cdot\mathbf{F}_c - \int_0^t \mathrm{d}s\ 2\mathsf{K}_0\beta_0\alpha(s) + \int_0^t \mathrm{d}s\ d\mathrm{N}\alpha(s) \\
&= -\int_0^t \mathrm{d}s\ \beta_0\mathbf{J}(s)\mathrm{V}(s)\cdot\mathbf{F}_c
\end{aligned} \tag{6.19}$$

which implies

$$\overline{\Omega}_t \equiv \frac{1}{t}\int_0^t \mathrm{d}s\ \Omega(\mathbf{\Gamma}(s)) = -\beta_0[\overline{\mathbf{JV}}]_t\cdot\mathbf{F}_c, \tag{6.20}$$

where

$$[\overline{\mathbf{JV}}]_t \equiv \frac{1}{t}\int_0^t \mathrm{d}s\ \mathbf{J}(s)\mathrm{V}(s). \tag{6.21}$$

Substituting this definition of the general dissipation function into the master Fluctuation Theorem equation (Equation (3.23)) gives

$$\ln\Big[\frac{\Pr(-\beta_0[\overline{\mathbf{JV}}]_t\mathbf{F}_c = \mathrm{A})}{\Pr(-\beta_0[\overline{\mathbf{JV}}]_t\mathbf{F}_c = -\mathrm{A})}\Big] = \mathrm{A}t. \tag{6.22}$$

As mentioned earlier, the use of a Nosé-Hoover thermostat and barostat will yield the same results. Indeed, any combination of Nosé-Hoover or Gaussian thermostat/barostat will yield the same expression for the isothermal-isobaric Fluctuation Theorem.

6.4 The SHREP Potential

In 2000, Hess and Kröger published a potential that could be used to model the repulsive part of interparticle interactions, instead of using WCA or the standard Lennard-Jones potential in molecular dynamics simulations [56]. The potential, denoted SHREP (**SH**ort range **REP**ulsive), was designed to overcome some of the shortcomings of WCA.

The mathematical form of the SHREP potential is

$$\Phi(\mathrm{q}) = \begin{cases} (9 - 8\mathrm{q})^3 & q < 1.125, \\ 0 & q \geq 1.125. \end{cases} \tag{6.23}$$

It has continuous 1st and 2nd derivatives for all q, in contrast to the more common truncated Lennard-Jones and WCA potentials. This allows for one to use larger time steps in MD simulations without the introduction of large numerical errors. Additionally, the SHREP potential has no singularity at $q = 0$ and has a cutoff that is smoother than WCA. In tests we performed in the isothermal-isobaric ensemble, we found that systems where the interaction potential was modelled via the SHREP potential could have a time step of up to an order of magnitude larger than systems using the Lennard-Jones potential truncated at $q = 2.5$ [55].

6.5 MD Simulations

Two MD simulations were performed that were designed to test Equation (6.22). The interaction potential used in the first simulation was Lennard-Jones (§2.3), while the second simulation used SHREP.

The conditions common to both simulations were N = 98, T = 1.0 and $n = 0.4$. For the simulation using the Lennard-Jones interaction potentials, the other conditions were $\mathsf{p}_0 = 0.968$, time step=0.0001. The simulation utilising SHREP had $\mathsf{p}_0 = 0.396$ and time step=0.001. The volume in the mathematical expressions presented in this chapter is the (varying) area of the unit cell in

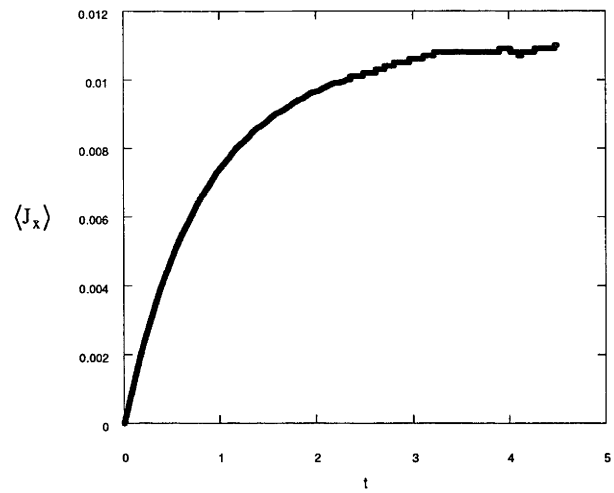


Figure 6.1: The change of the ensemble-averaged transient response of the colour current density over the length of a transient trajectory for a simulation utilising the SHREP potential in two dimensions with a Maxwell relaxation time of approximately $\mathcal{M}_T = 1$. The simulation time was $t = 4.5$, N = 98, $T_0 = 1.0$, $n = 0.4$, $p_0 = 0.396$, time step=0.001 and $F_c = 0.05$.

our two-dimensional MD simulations.

At time $t = 0$ the external colour field is applied and the response of the system is monitored for a time t, which is the trajectory length for the system. The simulations we performed tested the DFT (§3.4.1), TFT (§3.4.2) and the SSFT (§3.4.3).

6.6 Results

Figures 6.1 through 6.4 show the results of our MD simulations in the isothermal-isobaric ensemble that verify the FT. Figure 6.1 shows the response of the ensemble-averaged colour current density for an MD simulation using the SHREP potential for interparticle interactions. The Maxwell relaxation time ($\mathcal{M}_T$) for this system is approximately 1. The field strength in this simulation was $F_c = 1.0$.

Figure 6.2 shows the results of a TFT simulation with transient trajectory segments of length $t = 1.0$. Interparticle interactions were modelled by the Lennard-Jones potential with a cutoff radius of $r_{cut} = 3.0$. The time step used was 0.0001, in order for the Gaussian barostat to function properly, since the

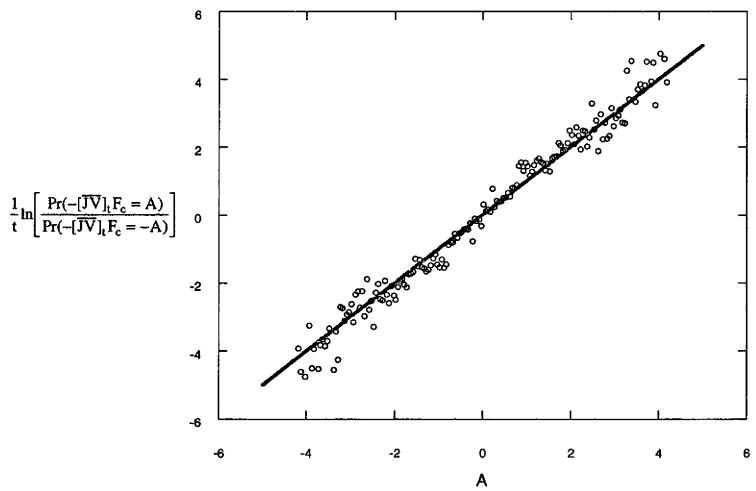


Figure 6.2: The results of a test of the TFT for a system in the isothermal-isobaric ensemble. The non-equilibrium transient trajectory segments were of length $t = 1.0$, which is approximately equal to the Maxwell relaxation time of the system. The interparticle interactions were modelled with the Lennard-Jones potential, using a cutoff radius of $r_{cut} = 3.0$. Other simulation conditions were N = 98, $T_0 = 1.0$, $n = 0.4$, time step=0.0001, $p_0 = 0.968$ and $F_c = 0.4$. The black circles show the results of the MD simulation, while the red line is the prediction made by applying Equation (6.22).

Lennard-Jones potential is much more sensitive to pressure fluctuations than the SHREP potential. The black circles are the MD simulation results and the red line is what is predicted by the isothermal-isobaric FT (Equation (6.22)).

The scatter of the MD data points in the plot is caused by insufficient statistics. If one allowed the simulation to run for a longer time, the data points would gradually converge to the predicted line. Notice that the points near the centre of the plot are very close to the predicted line. These points are the result of much more averaging (and correspond to points close to the centre of a probability histogram modelled by a Gaussian distribution) than the points farther from the centre.

Figure 6.3 shows a plot of the results of two TFT MD simulations. The data are from simulations with transient trajectory lengths of $t = 3.5$ (blue circles) and $t = 4.5$ (black crosses). Interparticle interactions have been modelled with the SHREP potential and a time step of 0.001 was used. The red line is the prediction of the isothermal-isobaric FT. Clearly the FT is verified in both cases, as expected.

The results of DFT and SSFT simulations are shown in Figure 6.4. Here the DFT results (blue circles) are obtained from a single steady-state non-equilibrium trajectory, which is divided into individual steady-state trajectories. The SSFT results (black crosses) are obtained from a simulation with an ensemble of transient trajectories with lengths of $t = 16$. The averaging of the data to test the SSFT is done from $t = 4$ (by which time the system has already reached a steady state) to $t = 16$. The red line is the prediction of the isothermal-isobaric FT. As we expect, both the DFT and SSFT are verified in this ensemble. Error bars are shown, which verify that the data points farther from the centre of the plot need longer simulation runs to improve the statistics before better agreement with the FT prediction will be observed.

6.7 Summary

We have shown that the Fluctuation Theorem is valid in the isothermal-isobaric ensemble. We have used Gaussian thermostat and barostat multipliers to generate the necessary isothermal-isobaric dynamics. We have also found that the SHREP potential is more stable in numerical simulations than is the Lennard-Jones potential. The Lennard-Jones potential can still be used to perform simulations in this ensemble, but the small time step necessary to avoid large numerical error is one order of magnitude less than that which may be used with the SHREP potential. This makes the SHREP potential much more suitable for such simulations.

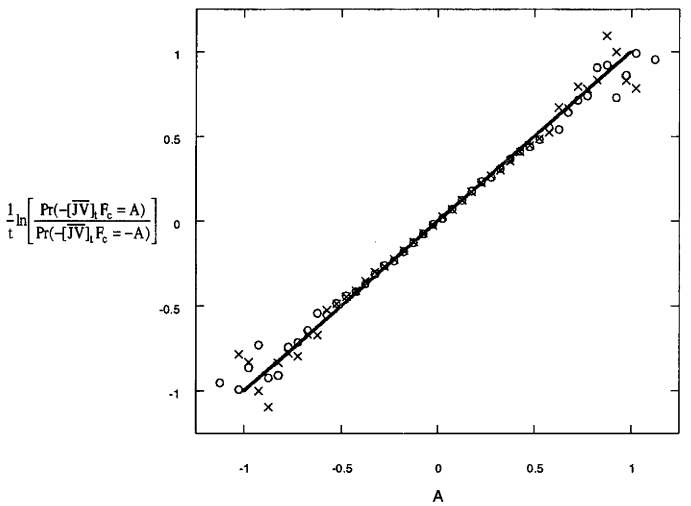


Figure 6.3: A plot of the results of a test of Equation (6.22) for two TFT MD simulations in two dimensions. The two sets of data displayed are the results of simulations with transient trajectory lengths of $t = 3.5$ (shown as blue circles) and $t = 4.5$ (shown as black crosses). Inter-particle interactions were modelled via the SHREP potential. Other simulation conditions were $T_0 = 1.0$, $n = 0.4$, $p_0 = 0.396$ and $F_c = 0.05$. The red line is that predicted by the isothermal-isobaric FT.

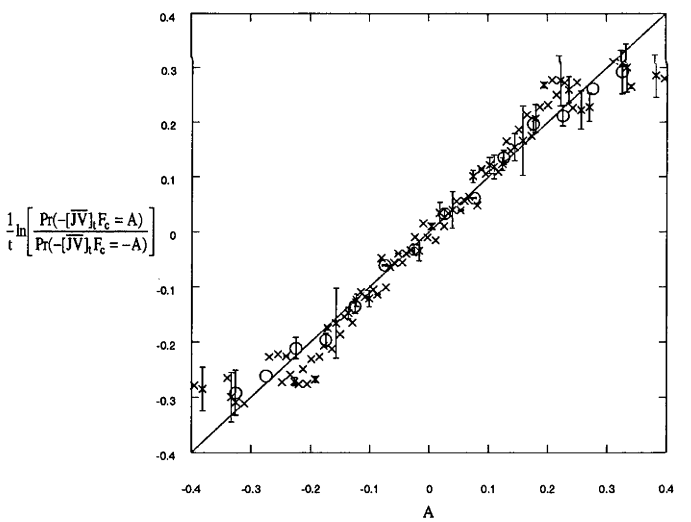


Figure 6.4: A plot of the result of tests of the SSFT (black crosses) and DFT (blue circles) given by Equation (6.22). The non-equilibrium side trajectories had length $t = 16$. The averaging for the SSFT was done from $t = 4.0$ to $t = 16$, as the system had reached a steady state before $t = 4.0$. The DFT trajectory segments were obtained from a single non-equilibrium trajectory in phase space that was divided into individual steady-state segments. Both simulations used the SHREP potential with N = 98, time step=0.001, $T_0 = 1.0$, $n = 0.4$, $p_0 = 0.396$ and $F_c = 0.05$. The red line in the plot is that predicted by the isothermal-isobaric FT (Equation (6.22)).

Chapter 7

Optical Tweezer Simulation and Wall-thermostatted FT

7.1 Introduction

Since the Fluctuation Theorem can be used to express the probability that the Second Law of Thermodynamics may be violated, it is perhaps a natural first reaction to assume that a laboratory experiment which would verify the FT would be impossible to perform, since we do not observe Second Law violating events in every-day life. There exist contemporary laboratory experiments, however, that are capable of effecting nano-scale measurements of picoNewton forces [57]. Measurements on such small scales of such small forces facilitate the testing of the Fluctuation Theorem via the observation of the microscopic negative fluctuations in the entropy of the system.

For machines built on such small (i.e. nano) scales, the Fluctuation Theorem becomes very important, as anti-events occur occasionally. The anti-events experienced by such small machines cause the machines to run in the reverse direction to that which is desired. For this reason, the FT defines a fundamental limit in nature: the smallest size at which a machine is able to produce useful work is the small-size limit for usable machines. Small components of biological systems, such as mitochondria, are of the size where the anti-events predicted by the FT affect the efficiency with which they function.

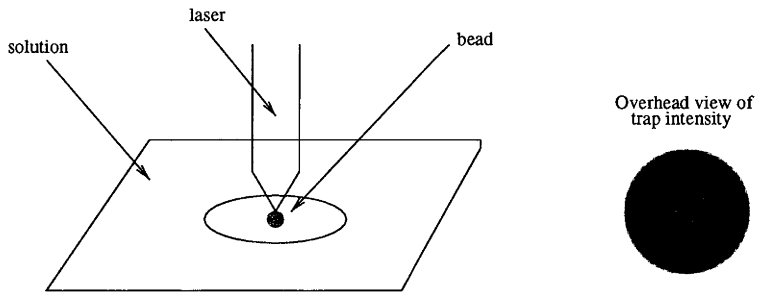


Figure 7.1: (left) Illustration of the optical tweezer apparatus used to test the Fluctuation Theorem. The optical trap (i.e. laser) is mounted on a movable stage. As it moves, it drags the bead through the solution. (right) The bead (shown in red) is trapped by the laser in a harmonic oscillating potential well. The region at the centre of the laser is the highest intensity, which decreases as one moves farther from the centre.

There are several possible ways in which a nano-scale experiment could test the FT. One such method is the trapping of a particle (or molecule) in a field and measuring the fluctuations in its instantaneous velocity. An exact description of such an experiment is detailed in the following section.

7.2 Experimental verification of the FT

This work is done in collaboration with Dr. Edith M. Sevick and Dr. Genmiao Wang.

The construction of a laboratory apparatus to test the FT involves an apparatus capable of monitoring small scale velocity fluctuations. The recent successful laboratory test [57], referred to in the introduction, involves the use of a laser mounted on a movable stage. The laser is used to trap a micron-sized transparent bead, which has an index of refraction that is greater than the solution in which it is placed. Figure 7.1 illustrates how this "optical tweezer" (OT) apparatus traps the bead in this experiment. The bead is drawn to the region of highest intensity of the laser beam, as the light from the laser is refracted with different intensity over the surface of the bead. The power of the laser is adjustable. The beam exerts a force on the bead, which is approximately harmonic.

The bead may be thought of as being trapped in a potential well, as depicted

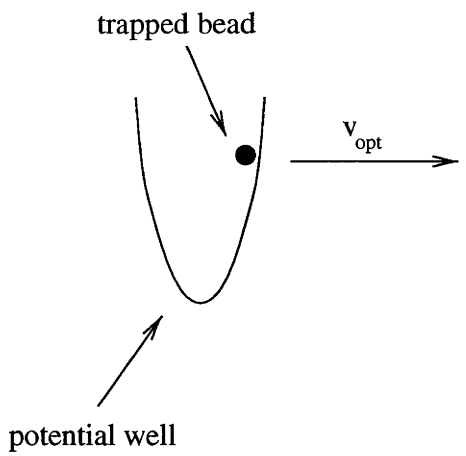


Figure 7.2: The potential well created by the laser of the optical trap as it holds the bead. As the optical trap translates on its stage, the bead moves while held by the laser, causing the bead to experience a force that varies in magnitude. The potential is harmonic and the bead oscillates in the harmonic well as shown in this figure. The potential well moves with a constant velocity $\mathbf{v}_{opt}$ as the stage upon which the laser is mounted moves.

in Figure 7.2. The mathematical form of the potential is

$$\Phi_{opt}(\mathbf{q}) = \frac{1}{2}k(\mathbf{q} - \mathbf{q}_0)^2, \tag{7.1}$$

where k is the force or trapping constant, x is the position of the bead and x_0 is the position of the centre of the optical trap. The force applied by the optical trap is

$$\mathbf{F}_{opt} \equiv -\frac{\partial \Phi_{opt}(\mathbf{q})}{\partial \mathbf{q}} = -k(\mathbf{q} - \mathbf{q}_0). \tag{7.2}$$

After the bead is trapped, the optical trap (i.e. the laser) begins to move, with velocity $\dot{\mathbf{q}}_0 \equiv \mathbf{v}_{opt}$, relative to the solution in which the bead sits. As it moves, the optical trap does work on the system equal to $\mathbf{v}_{opt} \cdot \mathbf{F}_{opt}(t)$.

Figure 7.3 depicts the situation of the bead as it is pulled through the solution by the optical trap. As the bead is pulled, it experiences a drag force from the surrounding solvent molecules. After a time a steady state is eventually reached, whereby the force from the optical trap and the drag force balance one another. Since the stage is driven by a piezo crystal it can only move so far. The limit of the stage movement is eventually reached and the

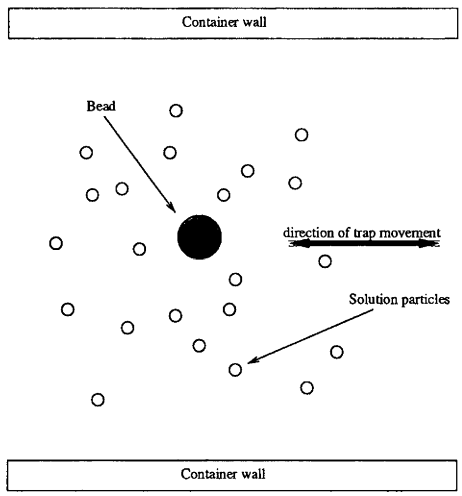


Figure 7.3: Diagram of the optical tweezer experiment. The bead is pulled through the solution by the optical trap as it moves upon its stage. The size of the bead is large compared to the particles in solution.

optical trap begins to move in the opposite direction.

In this experiment, the entropy production is

$$\overline{\Sigma}_t t = \beta \int_0^t \mathrm{d}s \; \mathbf{v}_{opt} \cdot \mathbf{F}_{opt}(s). \tag{7.3}$$

Note that no special mechanism is required to maintain the temperature of the solution at a constant value and the solution temperature is taken to be the same as the surroundings (i.e. the room in which the experiment is conducted). Equation (7.3) can be used to test the Fluctuation Theorem. Results of these experimental tests are shown later in this chapter.

7.3 MD Simulation of the OT Experiment

In order to simulate the laboratory experiment, it is necessary to find a method for emulating the optical trap. It is possible, as discussed in (§2.7), to divide the particles in an MD simulation into different groups that respond in different ways to an external field. In colour diffusion, the particles are normally divided equally into two groups. If we wish to examine how a single particle responds to the external field, we may set the colour field coupling coefficients, c_i, to zero for every particle except the one in which we are interested. This particle would

therefore emulate the bead trapped by the laser of the laboratory experiment and would be "trapped" by the external colour field that emulates the field generated by the laser.

Since the laboratory experiment begins with the laser trapping the bead at time $t \leq 0$ and then continuing to trap the bead for the duration of the experiment, the colour field must be active at all times when the simulation is running. In the laboratory experiment, the stage begins to move at time $t > 0$. An MD simulation of the OT experiment therefore needs to explicitly consider the movement of the stage in the equations of motion.

In order to model the laboratory experiment as accurately as possible, the top and bottom walls of the simulated box are at constant temperature. This is accomplished via the use of wall particles that have a thermostatting term in their equations of motion. Either a Gaussian or Nosé-Hoover thermostatting mechanism may be used in this regard. The particles in the wall must also have a restoring force applied to them to ensure that they remain in their original positions and do not drift into the solution.

The external field, which is applied to the trapped particle, must be harmonic in order to model the optical trap (Equation 7.1). The mathematical form of this potential is

$$\Phi_{opt}(\mathbf{q}) = \frac{1}{2} k \delta_{1i} (\mathbf{q}_i - \mathbf{q}_0)^2 \tag{7.4}$$

and the external colour field is therefore

$$\mathbf{F}_{opt} \equiv -\frac{\partial \Phi_{opt}(\mathbf{q})}{\partial \mathbf{q}_i} = -k \delta_{1i} (\mathbf{q}_i(t) - \mathbf{q}_0(t)) \tag{7.5}$$

where δ_{1i} is the Kronecker delta, k is the trapping constant, $\mathbf{q}_i(t)$ is the position of the trapped particle at time t and $\mathbf{q}_0(t)$ is the current position of the stage. Note that the Kronecker delta operates on one particle only. Equation (7.5) assumes that particle number one is the particle which will be trapped (i.e. the equivalent of the bead in the experiment). Therefore, in an MD simulation of this experiment, the equations of motion must include an extra equation to

simulate the movement of the particle with time, relevant to the position of the stage at a given time t. The equations of motion for this system are then

$$\begin{aligned}\dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - k\delta_{1i}(\mathbf{q}_i(t) - \mathbf{q}_0(t)) + \mathrm{S}_i\mathbf{F}_i^{\mathrm{w}} - \mathrm{S}_i\zeta\mathbf{p}_i \\ \dot{\mathbf{q}}_0(t) &= \mathbf{v}_{opt}\end{aligned} \tag{7.6}$$

here $\mathbf{v}_{opt}$ is the velocity of the stage, $\mathrm{N_w}$ is the number of particles in the thermostatted walls, ζ is the Nosé-Hoover thermostat multiplier (§2.9.3). $\mathrm{F}_i^{\mathrm{w}}$ is the restoring force for the wall particles, which has the form

$$\mathbf{F}_i^{\mathrm{w}} = -k_{wall}(\mathbf{q}_i - \mathbf{q}_i^{orig}), \tag{7.7}$$

where $\mathbf{q}_i^{orig}$ is the original lattice location of wall particle i. $\mathrm{S}_i = 1$ if the particle i is a thermostatting wall particle, $\mathrm{S}_i = 0$ otherwise. It may therefore be inferred that

$$\sum_i^{\mathrm{N}} \mathrm{S}_i = \mathrm{N_w} \tag{7.8}$$

is the number of wall particles in the system. The NH thermostat multiplier takes the form

$$\dot{\zeta} = \frac{1}{Q}\Big(\sum_{i=1}^{\mathrm{N_w}} \frac{\mathbf{p}_i^2}{m} - \mathrm{dN_W k_B T}\Big) \tag{7.9}$$

Although the simulations we have performed utilise a NH thermostatting mechanism, a Gaussian thermostat could also have been used. For this system, the Gaussian thermostat multiplier takes the form

$$\alpha = \Big(\frac{\sum_{i\in\mathrm{wall}}^{\mathrm{N_w}} \frac{\mathbf{F}_i^{\mathrm{w}}}{m}\cdot\mathbf{p}_i + \sum_{i\in\mathrm{wall}}^{\mathrm{N_w}}\sum_{j\neq i}^{\mathrm{N}} \frac{\mathbf{F}_{ij}}{m}\cdot\mathbf{p}_i}{\sum_{i\in\mathrm{wall}}^{\mathrm{N_w}} \mathbf{p}_i\cdot\mathbf{p}_i}\Big) \tag{7.10}$$

The use of a Gaussian thermostatting mechanism would imply a system characterised at equilibrium by an isokinetic distribution function. The form of the FT would, however, remain the same, as the FT is thermostat independent.

Note that in our derivation, as in the experiment, the laser can start at any

position (determined randomly) and does not always have to begin moving in the same direction.

7.4 Derivation of the Optical Tweezer FT

We now present a derivation of the Fluctuation Theorem [44] that can be used to verify the FT in the optical tweezer laboratory experiment and MD simulation. The derivation shown here assumes the system is in the canonical ensemble.

Consider a system where only wall particles are thermostatted and where only one particle responds to an external field. All of the particles are on a stage that moves at a constant velocity, $\mathbf{v}$, with time. For such a system, at time $t \leq 0$ the system is at equilibrium. The Hamiltonian for the system is

$$\mathcal{H}(\mathbf{\Gamma}, \mathbf{q}_0, t) = \mathcal{H}_0(\mathbf{\Gamma}) + \frac{1}{2}k(\mathbf{q}_1 - \mathbf{q}_0)^2 \tag{7.11}$$

where $\mathbf{q}_0$ is the initial position of the laser trap. The distribution function for this system is*

$$\begin{aligned} f(\mathbf{\Gamma}, \mathbf{q}_0, \zeta, 0) &\sim \mathrm{e}^{-\beta\mathcal{H}(\mathbf{\Gamma})+\frac{1}{2}Q\zeta^2} \\ &\sim \mathrm{e}^{-\beta\left[\mathsf{K}(\mathbf{p})+\Phi(\mathbf{q})+\frac{1}{2}k(\mathbf{q}_1-\mathbf{q}_0)^2\right]+\frac{1}{2}Q\zeta^2} \\ &\sim \mathrm{e}^{-\beta\left[\mathcal{H}_0(\mathbf{\Gamma})+\frac{1}{2}k(\mathbf{q}_1-\mathbf{q}_0)^2\right]+\frac{1}{2}Q\zeta^2} \end{aligned} \tag{7.12}$$

Using the definition of the general dissipation function (Equation (3.22)) we have

$$\begin{aligned} \overline{\Omega}_t t &= \int_0^t \mathrm{d}s\, \Omega(\mathbf{\Gamma}(s), \mathbf{q}_0(s)) \\ &= \ln\left(\frac{f(\mathbf{\Gamma}(0), \mathbf{q}_0(0), \zeta(0), 0)}{f(\mathbf{\Gamma}(t), \mathbf{q}_0(t), \zeta(t), 0)}\right) - \int_0^t \mathrm{d}s\, \Lambda(\mathbf{\Gamma}(s)) \end{aligned} \tag{7.13}$$

*see (§2.9.3)

For this system, the general dissipation function is then

$$\begin{aligned}\overline{\Omega}_t t &= \ln\Big(\frac{\mathrm{e}^{-\beta\mathcal{H}(\mathbf{\Gamma}(0),\mathbf{q}_0(0))+\frac{1}{2}Q\zeta^2}}{\mathrm{e}^{-\beta\mathcal{H}(\mathbf{\Gamma}(t),\mathbf{q}_0(t),\zeta(t))+\frac{1}{2}Q\zeta^2}}\Big) - \int_0^t \mathrm{d}s\ \Lambda(\mathbf{\Gamma}(s)) \\ &= \beta\big[\mathcal{H}(\mathbf{\Gamma}(t),\mathbf{q}_0(t)) + \frac{1}{2}Q\zeta^2(t) - \mathcal{H}(\mathbf{\Gamma}(0),\mathbf{q}_0(0)) + \frac{1}{2}Q\zeta^2(0)\big] - \int_0^t \mathrm{d}s\ \Lambda(\mathbf{\Gamma}(s)) \\ &= \int_0^t \mathrm{d}s\ \beta\dot{\mathcal{H}}(\mathbf{\Gamma}(s),\mathbf{q}_0(s)) + Q\zeta\dot{\zeta} - \int_0^t \mathrm{d}s\ \Lambda(\mathbf{\Gamma}(s))\end{aligned} \tag{7.14}$$

From the equations of motion we have

$$\dot{\mathcal{H}} = -k(\mathbf{q}_1 - \mathbf{q}_0)\cdot\mathbf{v}_s - \zeta\sum_i^{\mathrm{N}} \mathbf{p}_i\cdot\mathbf{p}_i\,\mathrm{S}_i \tag{7.15}$$

Combining equations (7.14) and (7.15) yields

$$\begin{aligned}\overline{\Omega}_t t &= -\beta k(\mathbf{q}_1 - \mathbf{q}_0)\cdot\mathbf{v}_{opt} \\ &\equiv \beta[\overline{\mathbf{F}_{opt}}]_t\cdot\mathbf{v}_{opt}\end{aligned} \tag{7.16}$$

which now allows us to write the expression for the Fluctuation Theorem for this system

$$\ln\Big[\frac{\Pr(\beta[\overline{\mathbf{F}_{opt}}]_t\cdot\mathbf{v}_{opt} = \mathrm{A})}{\Pr(\beta[\overline{\mathbf{F}_{opt}}]_t\cdot\mathbf{v}_{opt} = -\mathrm{A})}\Big] = \mathrm{A}t \tag{7.17}$$

7.5 MD simulations

In order to perform the simulations outlined in (§7.3) there are several considerations that must be made as to how one can model this system computationally. These include construction of the themostatted walls and storage problems associated with two different groups of particles (i.e. wall and solution particles).

7.5.1 Simulation of the thermostatted walls

In the experiment, there are walls on two opposite sides of the solution container. The container has only these two sides; the bead or any solution par-

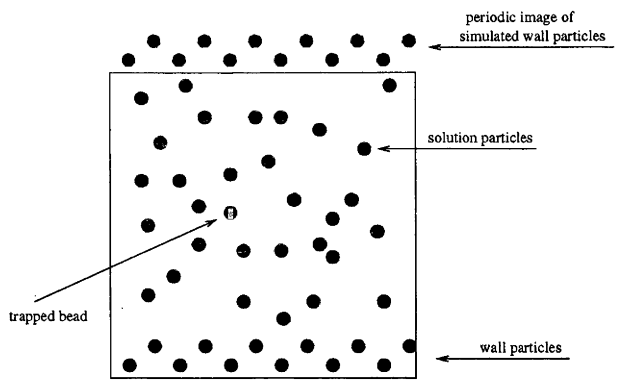


Figure 7.4: An illustration of the optical tweezer system with thermostatted walls. Two rows of particles are constrained from moving to form one wall. The wall on the opposite side of the box is created using periodic boundary conditions (§2.7.3), so there is no need to physically simulate a second wall of particles. In this example, the system has only 12 wall particles.

ticles are free to move in the other two directions. In order to simulate the thermostatted walls of the solution in the cheapest (in terms of computational time) and simplest (in terms of programming) manner, it is possible to simulate only one wall of particles and use periodic boundary conditions (§2.7.3) to replicate the other wall. The particles of the second wall are the periodic images of the particles in the first wall. These will interact with the solution particles and prevent them from leaving the box on that side. Figure 7.4 shows the arrangement of the simulated wall particles, solution particles and trapped particle in relation to one another.

In order to ensure that solution particles do not move through the walls, a minimum of two rows of wall particles should be used. The wall particles are held in place by the use of a restoring force in their equations of motion. Even if the restoring force is very large, the wall particles move small distances, but oscillate about their "proper" position. The movement of the wall particles, when arranged in two rows with sufficiently small spacing and with the proper restoring force constant, is never great enough to allow any solution particles to pass through.

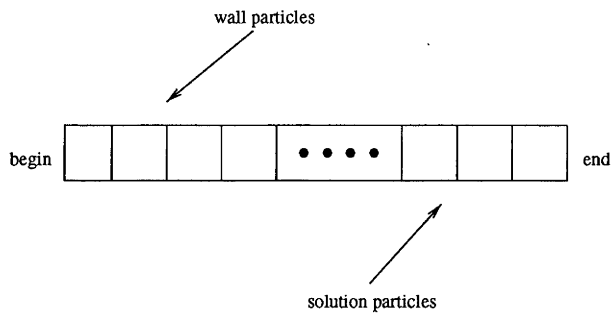


Figure 7.5: An array used to store the particles in an optical tweezer MD simulation. The array has two parts: the wall particles and solution particles. While there is no physical division between these particles in the array (that would necessitate a minimum of two arrays), the indices of the first wall and solution particles separate these two groups.

7.5.2 Computer Storage of Particles

Since there are effectively three types of particles in this MD simulation, an effective method of particle bookkeeping is needed. This may be accomplished in the simplest manner by using a single array to store all particles. Either the upper or lower end of this array is designated as being the wall particles and the other as the solution particles. Figure 7.5 gives a graphical depiction of one possible such array.

The trapped particle (simulated bead) need not be treated in a special manner, as a separate array will keep track of which particles (in this case only one) responds to the external colour field. Once the particles are arranged in the array, the index of the first wall particle and first solution particle are stored. When computing forces and integrating the equations of motion numerically, these indices will be used to determine how these operations are carried out (e.g. to know if the current particle should have terms in the equations of motion for a thermostat, if a wall particle).

7.6 Results

As mentioned in Chapter 3, the Fluctuation Theorem is more easily verified when many data are available. When a small number of statistics is available, the integrated FT (§3.8) is an easier form of the Fluctuation Theorem to verify.

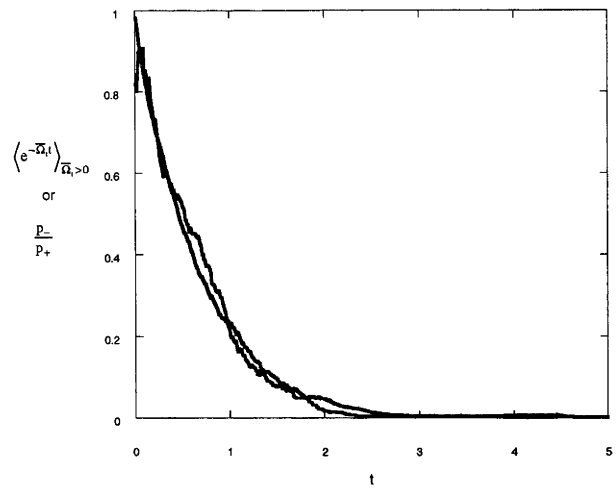


Figure 7.6: Results of a test of the integrated FT from the laboratory optical tweezer experiment. The blue line is the right-hand side of the IFT equation and the black line is the left-hand side. Ideally, the two curves should match. The data show that the laboratory experiment has verified the IFT.
Figure courtesy of Dr. Edith M. Sevick and Dr. Genmiao Wang

Therefore, the optical tweezer experiments test the IFT.

7.6.1 Experimental Results

The results of the application of the Fluctuation Theorem to the data obtained from the optical tweezer experiment are shown in Figures 7.6, 7.7, 7.8 and 7.9. Figure 7.6 shows the results of a test of the IFT. The mathematical form of the IFT (Equation (3.31)) is

$$\left[\frac{\Pr(\overline{\Omega}_t < 0)}{\Pr(\overline{\Omega}_t > 0)}\right] = \left\langle e^{-\overline{\Omega}_t t} \right\rangle_{\overline{\Omega}_t > 0}. \tag{7.18}$$

In the figure, the black curve shows the results of the left-hand side of this equation, while the blue curve shows the right-hand side. There is good agreement between the two curves, as is expected. The IFT is verified by these data.

Figures 7.7, 7.8 and 7.9 show the probability distributions for three different experiments. Each experiment used a different time over which the segments were averaged. The times used were $t = 10^{-3}s$, $t = 2s$ and $t = 3s$, respectively. From these three figures, it may be noted that as the averaging time

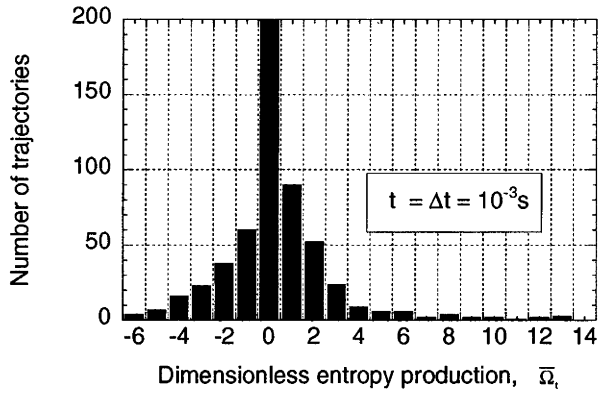


Figure 7.7: The probability histogram of data obtained from an optical tweezer laboratory experiment. The averaging time used was $t = 10^{-3}s$. The Second Law satisfying trajectories are those whose time-averaged irreversible entropy production appear to the right of the y-axis.
Figure courtesy of Dr. Edith M. Sevick and Dr. Genmiao Wang

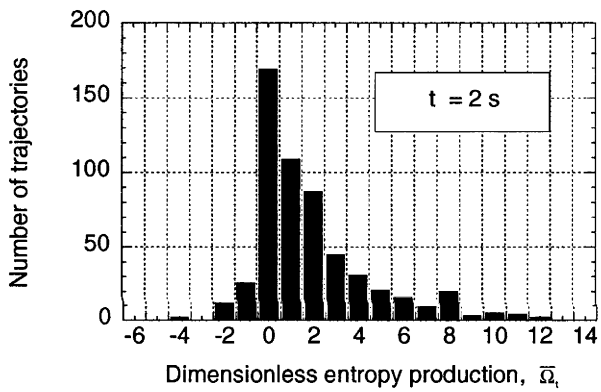


Figure 7.8: The probability histogram of data obtained from an optical tweezer laboratory experiment. The averaging time used was $t = 2s$.
Figure courtesy of Dr. Edith M. Sevick and Dr. Genmiao Wang

increases, the probability distribution shifts to the right. Once the averaging time has been increased to $t = 3s$, very few Second Law violating trajectories are observed.

7.6.2 MD Simulation Results

MD simulations were performed using the equations of motion and techniques outlined previously in this chapter. The conditions under which the simulations were run in reduced units (§2.7.4) were T=1.0, N=32, N_W = 24, n=0.3, time step (δ) = 0.001, $\mathbf{F}_{opt}$=1.0, v_s=0.5 and k_{wall} = 350.00. The walls of the system are thermostatted using the Nosé-Hoover mechanism. The initial equilibration

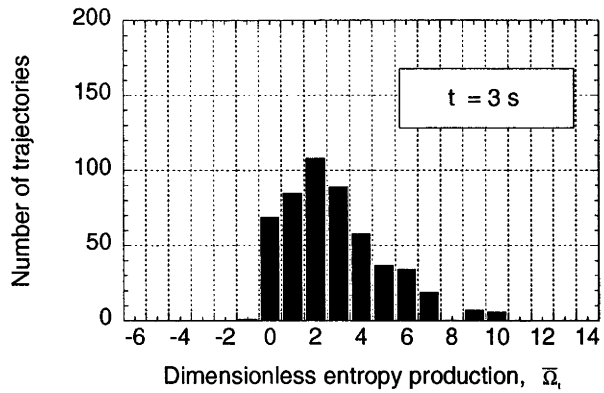


Figure 7.9: The probability histogram of data obtained from an optical tweezer laboratory experiment. The averaging time used was $t = 3s$. Note that there were no Second Law violating trajectories observed over the entire duration of the experiment, as all of the values for the irreversible entropy production shown in the histogram are positive.
Figure courtesy of Dr. Edith M. Sevick and Dr. Genmiao Wang

time for the simulation was $t_{equil} = 10^4$. During this time, the external field was applied to model the optical trap, but the velocity of the stage, $\mathbf{v}_{opt}$, was zero. At the end of each transient trajectory, the coordinates and momenta from the beginning of the trajectory were read in from a file. The coordinates and momenta were then used to further propagate the equilibrium trajectory for 200 time steps to avoid correlation between transient trajectory segments. The simulation performed in this manner never actually applies the time-reversal mapping, just as the experiment never actually attempts a reversal of the particle velocities. The MD simulation is therefore a close computational model of the laboratory experiment. The main differences are that the bead particle is identical to the solvent particles and the small number of particles in the simulation compared to in the experiment.

Figure 7.10 shows a plot of the results of an optical tweezer MD simulation that tests the IFT. The agreement between the left-hand (black circles) and right-hand (blue line) sides of the IFT equation is very good, thus demonstrating that the simulated system agrees with and verifies the IFT. For this simulation, the transient trajectory length was $t_{max} = 5000\delta = 5$. There were 343,500 trajectories for the averaging done in the interval $0 < t \leq 1$ and 33,400 in the range $1 < t \leq 5$.

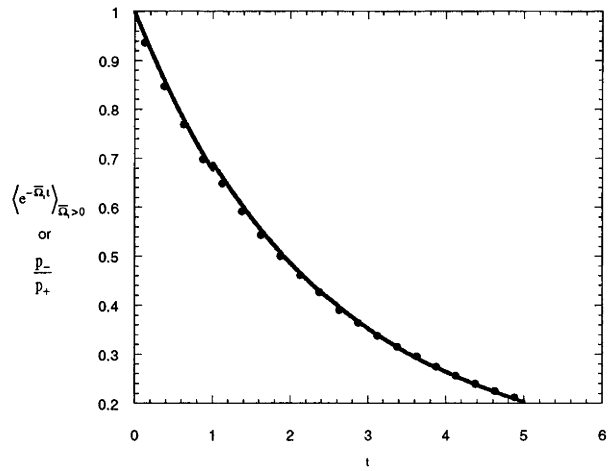


Figure 7.10: A plot of the results of an MD simulation that tests the IFT for 343,500 trajectories for $0 < t < 1$ and 33,800 trajectories for $1 < t < 5$. The blue line is composed of points calculated from the right-hand side of the IFT expression in Equation (7.18), while the black circles are the result of the evaluation of the left-hand side of the IFT expression. Near total agreement is observed, indicating that the MD simulation we have designed verifies the IFT.

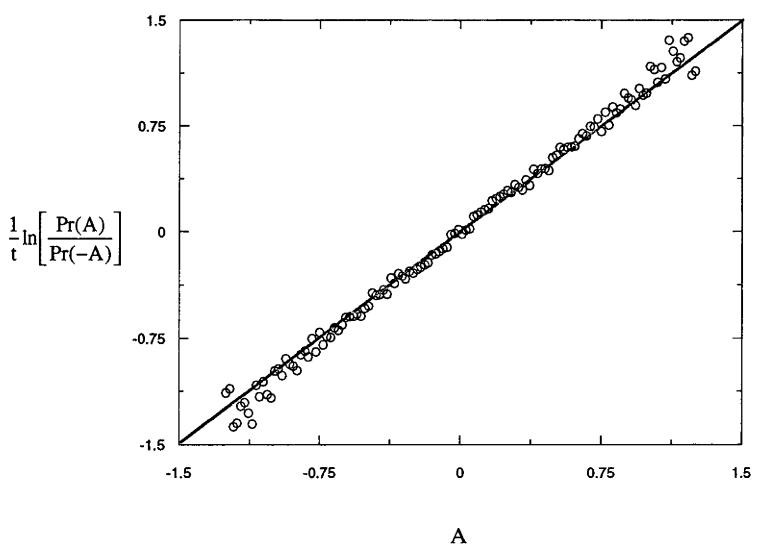


Figure 7.11: A plot of the results of an optical tweezer MD simulation that tests the TFT (Equation (7.17)). The numerical results of the simulation are shown as black circles, while a line of slope unity predicted by the FT is given in red. There is very good agreement between these two lines, indicating that the TFT is verified by this simulation.

Figure 7.11 shows a plot of the results of an optical tweezer MD simulation that tests the TFT. The results of the simulation are shown in circles, while a line with a slope of unity, predicted by the FT (Equation (7.17)), is shown in red. Ensemble averaging was done over 343,500 transient trajectories. Good agreement is noted between the slope of the line from the simulation results and that predicted by the FT. Note that the divergence at the ends of the simulated data from the prediction is caused by insufficient statistics. As the simulation is allowed to run for a longer period of time, the points that appear marginally divergent from the prediction, gradually approach the theoretical line. This is exactly what is predicted by the FT.

Chapter 8

The Time-dependent FT

8.1 Introduction

As mentioned in Chapter 3, the question of reversibility has been considered by scientists and philosophers for more than one hundred years. Boltzmann, Loschmidt and others such as Maxwell and the Ehrenfests spent much time considering why macroscopic systems in nature are irreversible. Loschmidt, who first published articles on reversibility and its implications on the Second Law of Thermodynamics, stated

> "Apparently, if the instantaneous velocities of all of the elements of any given system are reversed, the total course of the incidents must generally be reversed for every given system." [26]

On this subject, Boltzmann commented

> "In the whole universe, i.e., the totality of all individual worlds, events in the reversed direction indeed occur." [28]

Even though there was a consensus that the reversal of the equations of motion for a system is possible, contention remained as to why such anti-events are so rarely observed. With regards to this problem, Loschmidt remarked

> "The famous problem, the creation of anti-events from events, has no solution. Although, using simple instructions, the [solution] may

be put into words: reverse the instantaneous velocities of all of the atoms of the universe." [26]

It is interesting to note that the early scientists who were examining the mathematical reversibility of systems and its relationship to the Second Law seem not to have considered causality in their arguments. Loschmidt, Maxwell and Boltzmann refer to "events" and "anti-events" in the context of reversibility but never to causal or anti-causal events. However, if a process is observed in which cause precedes effect, then the time reversed anti-process must be one in which effect precedes cause. Such processes are then termed anti-causal.

8.2 Causality

8.2.1 The Axiom of Causality

It is possible to state, in mathematical terms, exactly what is meant by the terms causal and anti-causal [44, 48]. If we have an external force F that acts at a time t_2 for an infinitesimal time δt_2, the system will respond to this force. This response can be measured by defining a response function $\mathrm{L}(t_1, t_2)$. The component of the linear response ($\delta \mathrm{B}(t_1)$) is then

$$\delta \mathrm{B}(t_1) = \mathrm{L}(t_1, t_2)\mathrm{F}(t_2)\delta t_2. \tag{8.1}$$

If there is no relation between the time at which the experiment is performed and the response, then the response function, L, is dependent only on the difference between the time at which the force is applied and the time at which the response is monitored. We then have

$$\delta \mathrm{B}(t_1) = \mathrm{L}(t_1 - t_2)\mathrm{F}(t_2)\delta t_2. \tag{8.2}$$

This equation does not give any information about the state of the system. In order to define the causal and anti-causal response of a system, we must first distinguish between the components of the response that are causal or anti-causal. Causal response components would be those whose response depends

on forces of the past, while anti-causal response components would depend on forces from the future. The causal response component is then

$$\delta \mathrm{B}_C(t_1) \equiv +\mathrm{L}(t_1 - t_2)\mathrm{F}(t_2)\delta t_2, \qquad t_1 > t_2 \tag{8.3}$$

and the anti-causal component is

$$\delta \mathrm{B}_A(t_1) \equiv -\mathrm{L}(t_1 - t_2)\mathrm{F}(t_2)\delta t_2, \qquad t_1 < t_2. \tag{8.4}$$

We can then define the response due to the external field for all possible previous times, which defines the causal response

$$\mathrm{B}_C(t) = \int_{-\infty}^{t} \mathrm{L}_C(t - t_1)\mathrm{F}_c(t_1)\mathrm{d}t_1. \tag{8.5}$$

The anti-causal response is then the response due to the external field for all possible future times

$$\mathrm{B}_A(t) = -\int_{t}^{+\infty} \mathrm{L}_A(t - t_1)\mathrm{F}_c(t_1)\mathrm{d}t_1. \tag{8.6}$$

Equation (8.5) is also the **Axiom of Causality**, while Equation (8.6) is the **Axiom of Anti-causality**. The two are explicitly related via the response function, for it has been proven that [44, 48]

$$\mathrm{L}_C(t) = \mathrm{L}_A(-t). \tag{8.7}$$

The Fluctuation Theorem and Green-Kubo relations follow directly from the Axiom of Causality. Figure 8.1 shows the relationship between the Axiom of Causality, the Green-Kubo relations, the Fluctuation Theorem and irreversible entropy production. The corresponding Axiom of Anti-causality, anti-G-K relations, anti-FT and irreversible entropy consumption are shown on the left side of the figure. Note that the Green-Kubo relations for transport coefficients will have a definite sign (which is unknown until they are calculated). The anti-Green-Kubo relations will have a sign opposite to this.

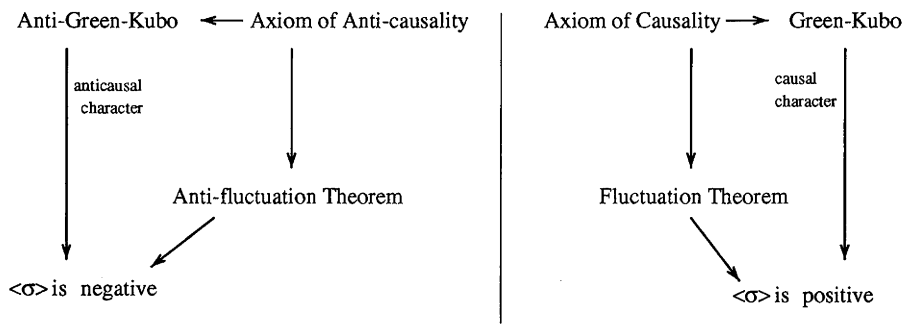


Figure 8.1: An illustration of the relationship between the Axiom of Causality, Fluctuation Theorem, Green-Kubo relations and irreversible entropy production (σ). The left side of the figure shows the equivalent anti-relationships.

It is important to note that there is no mathematical reason why we should prefer the Axiom of Causality over the Axiom of Anti-causality. The Axiom of Causality is taken to be the normal case, as it concurs with our daily observations as to the nature of the universe. Additionally, it would be difficult to imagine an anti-universe where the probability of present states could be predicted from the probabilities of future events. Boltzmann alluded to the assertion that if such an anti-universe existed, its occupants would not think such events to be unusual; rather, they would believe Second-law satisfying events to be exceptional.

8.2.2 Entropy and Causality

While the concepts of entropy and causality are related, the behaviour of one does not necessarily dictate the state of the other. A system could logically be entropy consuming, yet still be causal, or be entropy producing and anti-causal. *A priori* knowledge of the rate at which entropy is produced or consumed would not seem to be sufficient to determine the causal or anti-causal state of the system.

As an example, take the light source that can be activated by throwing a switch. If one were to throw this switch (the cause) one would see the light activated (the effect). This is an event with causal characteristics. If one then saw the same event in the reverse direction, two things would occur. The

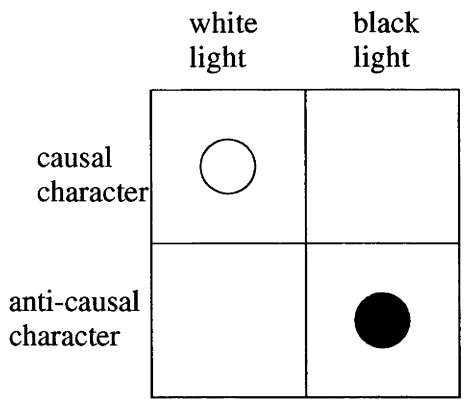


Figure 8.2: A matrix illustrating the relationship between sources of entropy production (white light) or consumption (black light) and the causal or anti-causal character of each. A white light can never have anti-casual character and a black light can never have causal character.

light would respond *before* the switch was thrown (an event that has anti-causal characteristics). Secondly, the light would absorb, rather than emit, photons. This is an event that is Second Law violating, as the disordered photons converge to a more ordered state; the entropy of the system decreases. To a direct observer of this sequence of anti-events, the light would appear black, since photons are being absorbed, rather than emitted. This "black light" also has distinctly anti-causal characteristics.* The conjugate event with causal character involves an increase of the entropy of the system after the switch is thrown and the photons move away from the light source.

Boltzmann gave some consideration to the anti-events that would be produced by fluctuations in the entropy of the universe. He believed that what we are able to observe of the universe is only a minute part; that the positions and velocities of all of the atoms in our perceived universe comprise only one microstate. Boltzmann stated that the entropy reducing states required for the formation of intelligent life is the result of a fluctuation in the entropy of a system comprising the entire universe. In most other areas of the universe, thermal equilibrium has already been reached and such fluctuations are no longer possible. He argued that although the massive fluctuation in the uni-

*The FT shows that it is unlikely to have black light that is causal in character, e.g. a black light that absorbs photons just after the switch for the light is thrown. Figure 8.2 illustrates this.

verse's entropy that has made life possible as we know it may appear to have nearly infinite duration to us, it is only an instantaneous fluctuation on the time scale of the universe. [28] With regards to this, Boltzmann wrote

> "Then there must be in the universe, otherwise everywhere in thermal equilibrium, i.e., dead, relatively small regions here and there with the size of our star space (let us call them individual worlds), that depart markedly from thermal equilibrium during the relatively short time of eons." [28]

8.3 Time-dependent Fields and Reversibility

In order to determine if a phase function responds before or after a change to the perturbing field takes place, it is necessary to use time-dependent external fields, rather than external fields that are constant at all times. Since in previous chapters the Fluctuation Theorem has dealt only with non-equilibrium systems that are forced from equilibrium by constant external fields, the applicability of the FT to these new time-dependent systems must be examined.[†] In order for a Fluctuation Theorem to be applicable for systems that have time-dependent external fields, there are some considerations that must be made, which were inconsequential in previous versions of the FT with constant external fields.

The derivation of the FT, outlined in (§3.5.4), applies to systems with time-dependent external fields only if several conditions are met.[‡] First, for every trajectory there must be a conjugate trajectory for which the time average of the dissipation functions are related by

$$\overline{\Omega}_t = -\overline{\Omega}^*_t. \tag{8.8}$$

We show that this means that the external field must have a definite parity

[†]Previously, a version of the FT was derived to which F_e and $\mathrm{M}^{\mathrm{T}}(\mathrm{F}_e)$ was applied (see Equation (31) of Reference [48]). This differs from the current investigation, as here the time-reversal mapping is never explicitly applied.

[‡]Here we assume that $\mathrm{F}_e(t)$ is the same for every trajectory in contrast to Reference [48].

under time-reversal symmetry. The parity of the external field refers to it being even or odd under time-reversal symmetry about t_{max} (as depicted in Figure 8.6). We can express this as

$$\mathrm{M}^{\mathrm{T}}[\mathrm{F}_e(t)] = \pm \mathrm{F}_e(t_{max} - t). \tag{8.9}$$

Additionally, to simplify the form of the FT, the initial distribution must be even under time-reversal,[§] i.e.

$$f(\mathbf{\Gamma}, 0) = f(\mathrm{M}^{\mathrm{T}}(\mathbf{\Gamma}), 0). \tag{8.10}$$

The final requirement is that the initial distribution and subsequent dynamics must be ergodically consistent, i.e.

$$f(\mathrm{M}^{\mathrm{T}}(\mathbf{\Gamma}(t)), 0) \neq 0 \; \forall \mathbf{\Gamma}. \tag{8.11}$$

Ergodic consistency ensures that the initial ensemble contains both the forward and time-reversed phases of the endpoints of all possible trajectories.[¶] If these conditions are met for a system with a time-dependent external field, then the derivation of the Fluctuation Theorem given in (§3.5.4) is also valid in this case.

8.3.1 Numerical Reversibility and Fluctuating Fields

In order to ascertain the degree to which a system (whose dynamics is determined numerically) is reversible, we can test the system with the proposed external field and conditions. To initiate such a test, we propagate a single trajectory in the forward direction. After a time t_{max}, a time-reversal mapping is applied,[‖] and the anti-trajectory is propagated forward for the same amount

[§]Note that if Equation (8.10) is not true, a FT can still be obtained, but it will look more complicated.

[¶]Note that if a system is not ergodically consistent, then the FT cannot be applied, even in cases where the external field is constant. An example of ergodic inconsistency would be a system with an initial distribution is in the microcanonical ensemble, but with dynamics that did not preserve the energy of the system.

[‖]as defined in (§3.5.2) and such that $\mathrm{F}_e(t)\mathrm{J}_{fwd}(t) = -\mathrm{F}_e(0)\mathrm{J}_{anti}(0)$

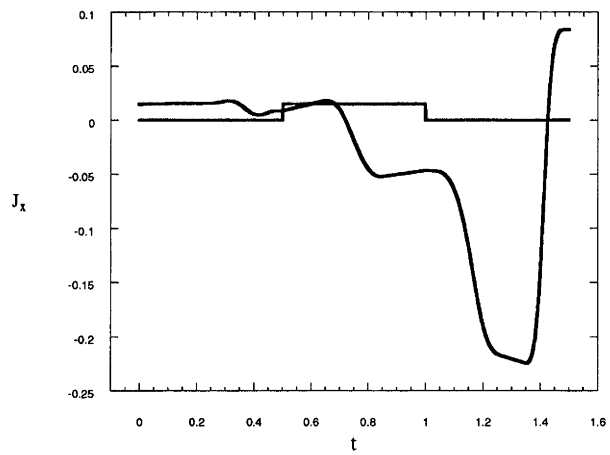


Figure 8.3: A plot of the response of the current (black line) from an MD simulation for an individual trajectory in the forward direction. The waveform of the external field is shown in red.

of time. If the system is indeed numerically reversible, then the response of the current to the external field for the anti-trajectory will be related (via the time-reversal mapping) to the response of the current for the forward trajectory. If the system is even or odd under the time-reversal mapping, then the anti-response of the current (J_{anti}) is related to the forward response (J_{fwd}) by

$$J_{anti}(t)F_e(t) = -J_{fwd}(t_{max} - t)F_e(t_{max} - t). \tag{8.12}$$

An MD simulation designed to test the reversibility of one test system was carried out. The conditions for the simulation were T=1.0, N=8, n=0.4, time step=0.001, $t_{max} = 1.5$, Nosé-Hoover thermostatting mechanism (§2.9.3) and a WCA potential (§2.3.1). The fourth-order Runge-Kutta algorithm was used for numerical integration. Figure 8.3 shows the response of the current for a system that uses a three-step time-dependent external field, which is even under the time-reversal mapping as discussed in (§8.3.2). Figure 8.4 shows the response of the current for the time-mapped trajectory. Note that the trace of the current in the Figure 8.4 is the time-reversal mapped conjugate of that in Figure 8.3. A plot of both of these curves together is shown in Figure 8.5. Here the current trace of the anti-trajectory has had a time-reversal map (M^T) applied to be more easily comparable to the forward trajectory by visual examination. The forward current trace is shown as black circles, while the

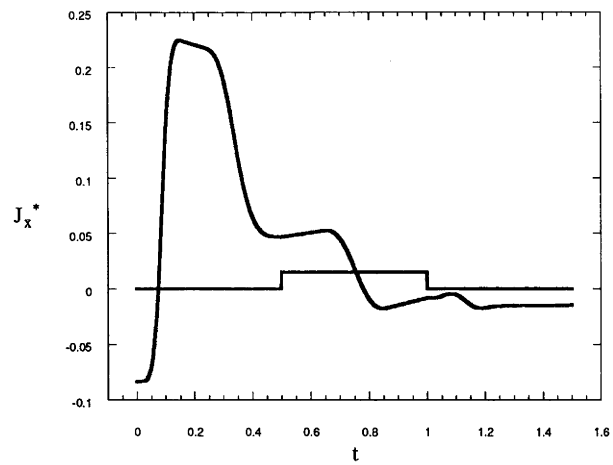


Figure 8.4: A plot of the response of the current from an MD simulation for a trajectory that is conjugate to that shown in Figure 8.3. Note that this trajectory was determined by numerically solving the equations of motion starting from a point $\mathbf{\Gamma}^*(0)$ that is the time-reversal map of the phase point at the end of the trajectory in Figure 8.3 (i.e. $\mathrm{M}^{\mathrm{T}}(\mathbf{\Gamma}(t))$). The waveform of the external field is shown in red.

mapped anti-current trace is shown as a blue line. Only every 20^{th} point of the forward curve is shown in order to facilitate comparison of the two curves. There is almost no difference in the two curves when viewed in this manner, verifying that for this field under these conditions, the numerical solution of the equations of motion for this system with this external field is time reversible.

If we examine the effects on the microscopic reversibility of a system of an external field that has parity neither even nor odd under the time-reversal mapping (§8.3.2), then we can have a better idea of exactly what effect such external fields have on the numerical reversibility of a system. We have taken such a system and propagated a single forward trajectory, applied the time-reversal mapping and propagated the system in the forward direction again. As with the three-step potential, the forward response matches the time-reversed forward response. These tests verify that the conditions we use for our numerical investigations of the time-dependent FT produce systems that are numerically reversible.

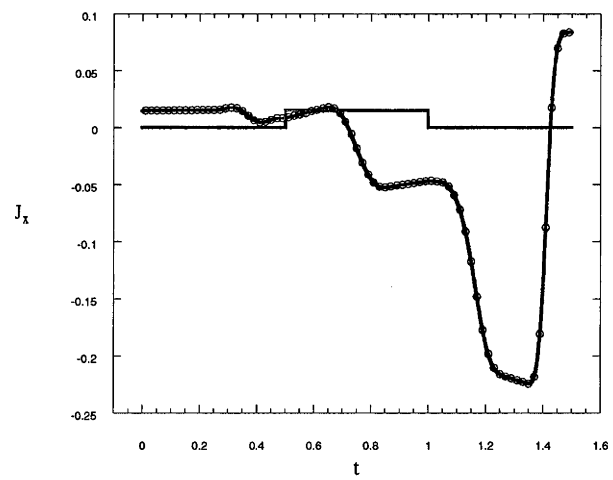


Figure 8.5: A plot showing both of the current responses from Figures 8.3 and 8.4. The response for the anti-trajectory (blue line) has been inverted and flipped so as to make the two curves visually comparable. It is superimposed over the response for the forward trajectory (black circles). Notice that the agreement between the two curves is almost absolute. It is expected that if a system is reversible, two such current traces should match exactly. Numerical error is responsible for any minute deviations between the two curves shown. Only every 20^{th} point in the forward curve is shown in order to facilitate visual comparison of the curves.

8.3.2 Step Fields

Step fields comprise one type of external field that can be used for MD simulations where a time-dependent external field is desired. As mentioned previously, for the derivation of the FT to apply to a system with a time-dependent field, the field must have even or odd parity under the time-reversal mapping (§3.5.2). If a time-dependent external field has no parity, then the time-reversed conjugate trajectory might not exist.

Figure 8.6 gives examples of time-dependent external fields that have different parities under the time-reversal mapping. The response of the current to the external field is shown, both before and after the time-reversal mapping is applied. For the field with no parity, the response after the application of the time-reversal mapping has no simple symmetry relation to the response before

the mapping, i.e.

$$\begin{aligned} \overline{\mathrm{J}}_{t_{max}} &\neq \overline{\mathrm{J}}^*_{t_{max}} \\ [\overline{\mathrm{JF}}_e]_{t_{max}} &\neq [\overline{\mathrm{JF}}^*_e]_{t_{max}}, \end{aligned} \tag{8.13}$$

where the asterisk indicates the time-reversal mapping has been applied. For the field with odd parity, there is a clear relationship between the response of the two trajectories

$$\begin{aligned} \overline{\mathrm{J}}_{t_{max}} &= \overline{\mathrm{J}}^*_{t_{max}} \\ [\overline{\mathrm{JF}}_e]_{t_{max}} &= -[\overline{\mathrm{JF}}^*_e]_{t_{max}}. \end{aligned} \tag{8.14}$$

For a field with even parity the response before and after the application of the time-reversal mapping is related by

$$\begin{aligned} \overline{\mathrm{J}}_{t_{max}} &= -\overline{\mathrm{J}}^*_{t_{max}} \\ [\overline{\mathrm{JF}}_e]_{t_{max}} &= -[\overline{\mathrm{JF}}^*_e]_{t_{max}}. \end{aligned} \tag{8.15}$$

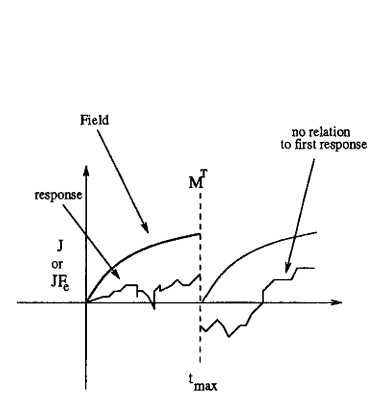
Note that the field is not reversed when the time-reversal mapping is applied.

Fields with odd or even parity

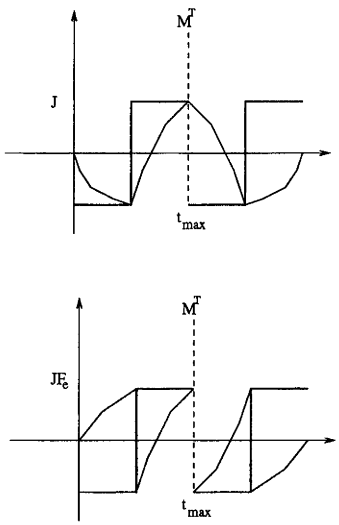
Examples of this group of time-dependent fields include three-step and two-step fields that allow the system to return to its original position in phase space upon application of the time-reversal mapping and propagation forward for the same time as before the time-reversal mapping was applied. Figure 8.7 illustrates a three-step field that has even parity under the time-reversal mapping, while Figure 8.8 depicts a step field that has odd parity under the time-reversal mapping.

Fields with neither even nor odd parity

This group of time-dependent external fields includes any field that, upon application of the time-reversal mapping to the field, has parity that is neither



Field with parity neither even nor odd



Field with odd parity

Figure 8.6: An example of an external step field that has parity neither even nor odd and one that has odd parity. The form of the external field is shown in red, along with the response of the system to it in black. t_{max} is the length of a trajectory segment over which the response is monitored. The dotted line indicates when the time-reversal mapping ($\mathrm{M^T}$) is applied. Note that only the field with odd parity allows the system to return to its starting position in phase space, since for the field with odd parity we have $\overline{\mathrm{J}}_{t_{max}} = \overline{\mathrm{J}}^*_{t_{max}}$ and $[\overline{\mathrm{JF}}_e]_{t_{max}} = -[\overline{\mathrm{JF}}^*_e]_{t_{max}}$. For the field with no parity we have $\overline{\mathrm{J}}_{t_{max}} \neq \overline{\mathrm{J}}^*_{t_{max}}$ and $[\overline{\mathrm{JF}}_e]_{t_{max}} \neq [\overline{\mathrm{JF}}^*_e]_{t_{max}}$. Here * indicates that the time-reversal mapping has been applied.

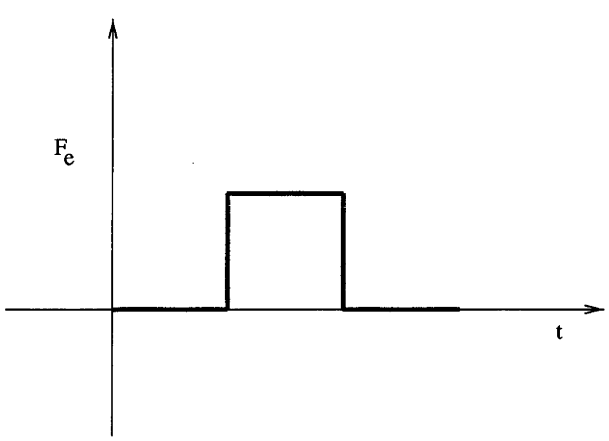


Figure 8.7: A three-step field that has even parity under the time-reversal mapping. The waveform of the field is such that the FT is applicable to systems that use this field.

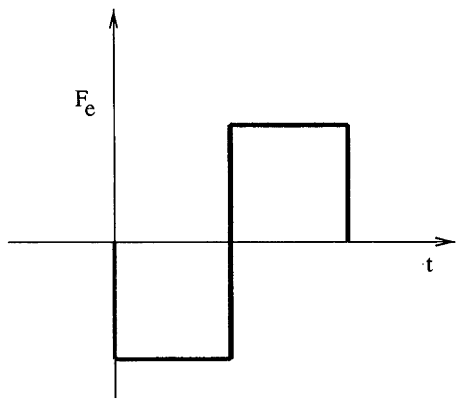


Figure 8.8: A two-step field that has odd parity under the time-reversal mapping. The waveform of the field is chosen so as to make the field odd under the time-reversal mapping (§3.5.2).

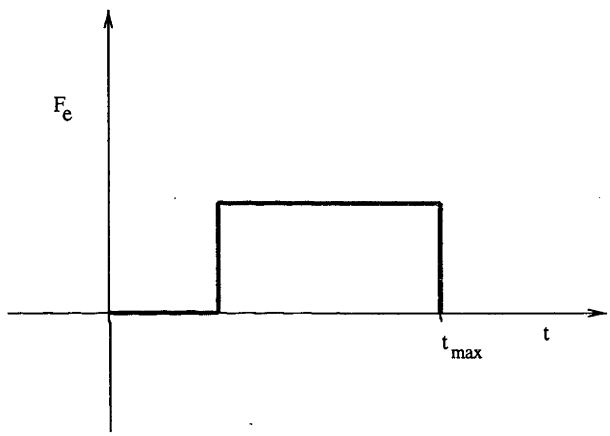


Figure 8.9: An example of a two-step external field that has neither even nor parity. The system will not be able to return to its original coordinates in phase space after application of the time-reversal mapping and propagation of the system for a time equal to that before the application of the mapping.

even nor odd. Figure 8.9 depicts an example of one such external field. Since the parity is neither even nor odd, the system will not be able to return to its original coordinates in phase space after the time-reversal mapping is applied** and the system is allowed to propagate for a time equal to that before the application of the mapping.

**Assuming the time-dependence of the field is not changed.

8.3.3 The Time-dependent FT

For the case of thermostatted systems with time-dependent external fields, we have a general dissipation function (§3.22) of the form

$$\begin{aligned}\overline{\Omega}_t t &= -\beta_0 \int_0^t \mathrm{d}s\ \mathrm{J}(s)\mathrm{F}_e(s)\mathrm{V} \\ &= -\beta_0 \overline{\mathrm{J}(t)\mathrm{F}_e(t)}\mathrm{V}.\end{aligned} \tag{8.16}$$

Applying the master Fluctuation Theorem equation (Equation (3.23))

$$\ln\left[\frac{\Pr(\overline{\Omega}_t = \mathrm{A})}{\Pr(\overline{\Omega}_t = -\mathrm{A})}\right] = \mathrm{A}t \tag{8.17}$$

to this problem with $\overline{\Omega}_t t$ as defined above yields

$$\ln\left[\frac{\Pr(-\beta_0 \overline{\mathrm{J}(t)\mathrm{F}_e(t)}V = \mathrm{A})}{\Pr(-\beta_0 \overline{\mathrm{J}(t)\mathrm{F}_e(t)}V = -\mathrm{A})}\right] = \mathrm{A}t. \tag{8.18}$$

8.4 MD Simulations

Molecular dynamics simulations were performed to investigate the effects of the external fields with either even/odd parity or no parity mentioned earlier. Two sets of tests were performed. The first was a test of the effects of reversibility on conjugate currents and the second was a direct test of the FT using Equation (8.18).

The equations of motion for these simulations are

$$\begin{aligned}\dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \mathbf{F}_c(\phi) - \zeta \mathbf{p}_i \\ \dot{\phi} &= \omega,\end{aligned} \tag{8.19}$$

where ζ is the Nosé-Hoover thermostat multiplier (§2.9.3), ω is the frequency of the periodic external field, $\phi(t)$ is a periodic function and $\mathrm{F}_c(\phi)$ is the external colour field.

In the following discussions of MD simulation results, unless stated otherwise, the conditions for the simulations are T=1.0, N=8, n=0.4, time step=0.001, $t_{max} = 1.5$, $F_c = 0.15$ with a WCA potential (§2.3.1).

8.4.1 Conjugate Currents

By histograming the responses on the basis of the time-averaged entropy production, we are able to directly compare the character of the response of the phase function in question. We divide the area under the probability distribution function for the dissipation function to the right of the y-axis into bins. Similarly, the area to the left of the y-axis is divided into bins of equal width to those created to the right. It is important that the width of each bin be equal, so as to be able to compare the response of a phase function as one moves farther to the right or left.

Consider a probability histogram of the time average of the dissipation function.[††] In such a histogram, we can examine the response of a phase function in a bin and its conjugate bin. Second Law satisfying transient trajectories will be those to the right of the y-axis. The Axiom of Causality dictates that these trajectories will be more probable than Second-law violating trajectories.

If the dissipation function changes before the field actually changes, then we would say the transient trajectory has anti-causal character. The opposite possibility, where the dissipation function changes after a change in the external field, which is, of course, what we normally expect, would be said to exhibit causal characteristics.

External field is even under time reversal

As an example of a probability histogram where we have binned the data presented in it, take Figure 8.10. This histogram was created using MD simulations to generate trajectories using a field that has even parity ($F_e = 0$, $t < 0.5$; $F_e = 0.15$, $0.5 < t < 1.0$; $F_e = 0$, $1.0 < t < 1.5$). Here there are seven bins (numbered 1 through 7) where the integrated entropy production

[††]That is positive according to the Second Law of Thermodynamics.

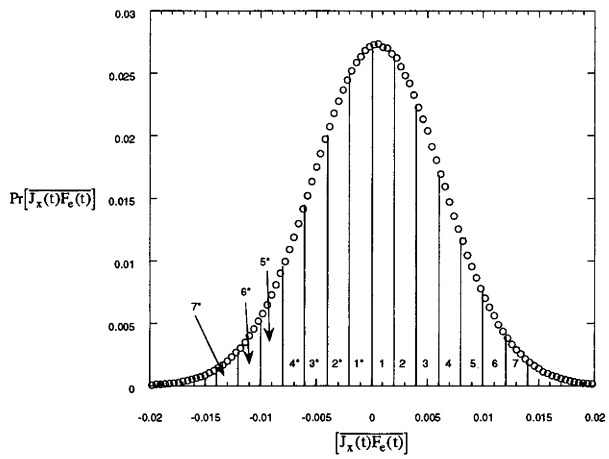


Figure 8.10: A probability histogram of the dissipative flux from an MD simulation, which has been partitioned into 14 bins. The seven bins to the right of the y-axis are the bins holding the Second-law satisfying responses of the dissipative flux to the external field. The seven bins to the left are the conjugate anti-bins holding the anti-currents of each of the corresponding bins.

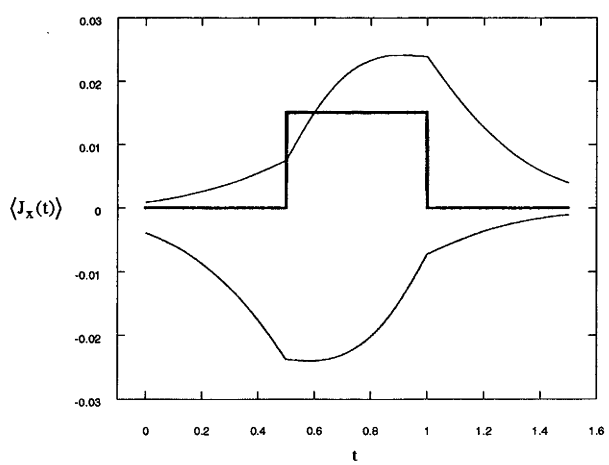


Figure 8.11: The sub-ensemble-averaged dissipative flux for bin 1 of Figure 8.10 is shown in black. The response in the conjugate anti-bin is shown in blue. The two response curves are related via the time-reversal mapping. The external field is shown in red.

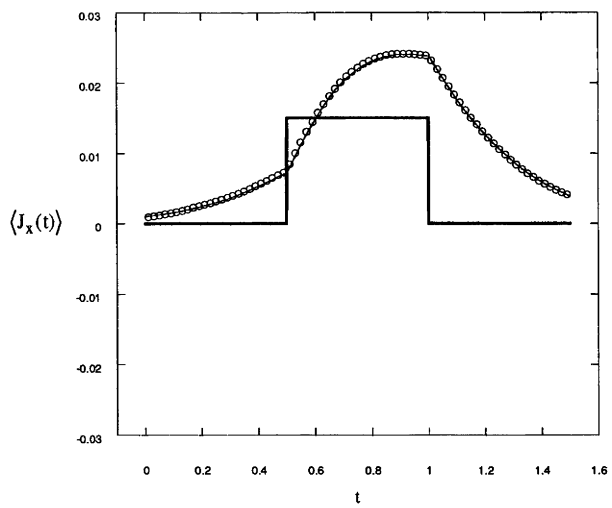


Figure 8.12: A plot with the sub-ensemble-averaged dissipative flux shown in Figure 8.11. The Second-law satisfying response is shown in black. Here the response that violates the Second Law (shown in blue) has been time-reversal mapped in order to facilitate the comparison of responses. The agreement between the two response curves is very good, indicating that this system with this external field is reversible. The field is divided by ten for purposes of the plot.

is positive and seven anti-bins (numbered 1* through 7*) containing responses which violate the Second Law. A plot of the dissipative flux in bins 1 and 1* along with the waveform of the external step field is shown in Figure 8.11. The response in the bin 1 is shown in black, while the conjugate anti-response (in bin 1*) is shown in blue. The two response curves are related via the time-reversal mapping. The same plot, with the time reversal map of the observed anti-response, shown in blue, flipped and inverted so as to make it more visibly comparable to the response, is shown in Figure 8.12.

Figure 8.13 depicts the response to the external field for bins 4 and 4*. Again, it can be seen that the response in a bin is directly related to the response in a conjugate bin via the time-reversal mapping. If we examine the response for all 7 bins (numbers 1 through 7 in Figure 8.10), as shown in Figure 8.14, we can see that the amplitude of the response increases as the time-integrated entropy production increases.

From Figures 8.11, 8.13 and 8.14 we see that in general, the bin-averaged responses are mixed with regard to their causal/anti-causal behaviour. In Figure 8.13 we can see periods where each of the responses appears to pre-empt

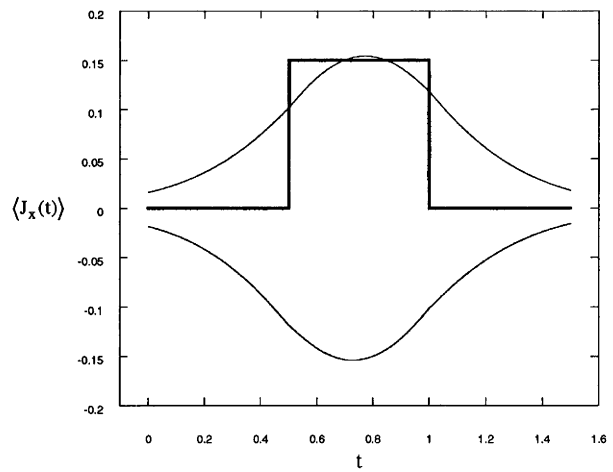


Figure 8.13: The Second-law satisfying and Second-law violating response of bins 4 and 4* to the external field (shown in red). As noted in previous figures, the response and conjugate response are related via the time-reversal mapping.

the changes in $F_c(t)$. Furthermore, Figure 8.14 illustrates the effect of choosing bins that are farther to the right of the y-axis. The anti-causal components of the response increases as the bins become more strongly entropy producing.

The response of the ensemble-averaged dissipative flux over *all* bins is shown in black in Figure 8.15. The external field is depicted in red. Here the response is initially zero until the external step field becomes positive, *after* which time the response follows the field. Note that the response is *entirely* causal in character, as it never pre-empts a change in the external field. In contrast to the sub-ensemble-averaged currents, the full ensemble-averaged current behaves in a normal "causal" manner.

The full ensemble-averaged dissipative flux shown in Figure 8.15 is the sum of the product of the ensemble-averaged dissipative flux in each bin and the weight of the response for that bin. Mathematically, we can express this as

$$< \mathrm{J}(t) >= \sum_{i}^{bins} \mathsf{w}_i < \mathrm{J}(t) >_i, \tag{8.20}$$

where w_i is the relative frequency of bin i. We know that the ensemble-averaged response for a single bin is the time-reversal mapping of the response in the conjugate bin, i.e.

$$< \mathrm{J}(t) >_i = \mathrm{M}^{\mathrm{T}} < \mathrm{J}(t) >_{i^*} . \tag{8.21}$$

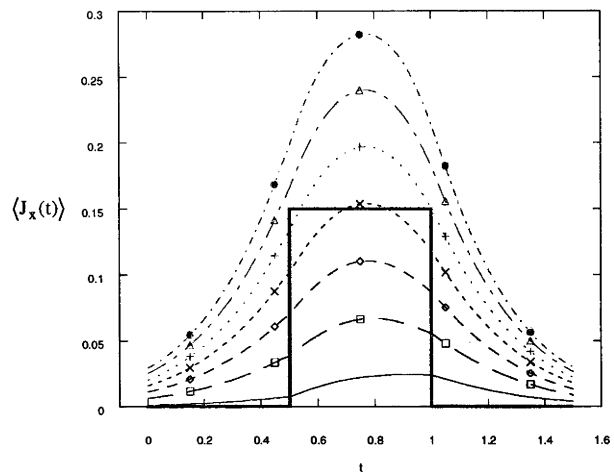


Figure 8.14: A plot of the dissipative flux of each of bins 1 through 7, as shown in Figure 8.10. As the time-integrated entropy production increases, the amplitude of the response increases. The external field is shown in red. The bins and their corresponding colours are 1 (black), 2 (blue with squares), 3 (green with diamonds), 4 (black with crosses), 5 (pink with vertical crosses), 6 (cyan with triangles) and 7 (orange with solid circles).

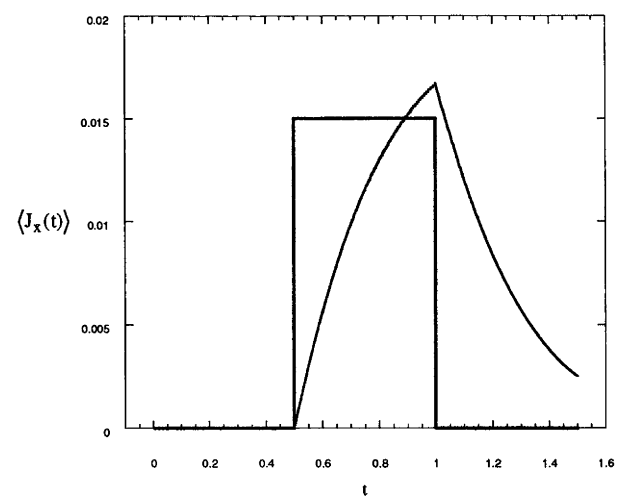


Figure 8.15: The ensemble-averaged dissipative flux for *all* bins (shown in black). As expected, the response is entirely causal in character, as it changes only after the external field changes. The external field is shown in red and is divided by ten for purposes of the plot.

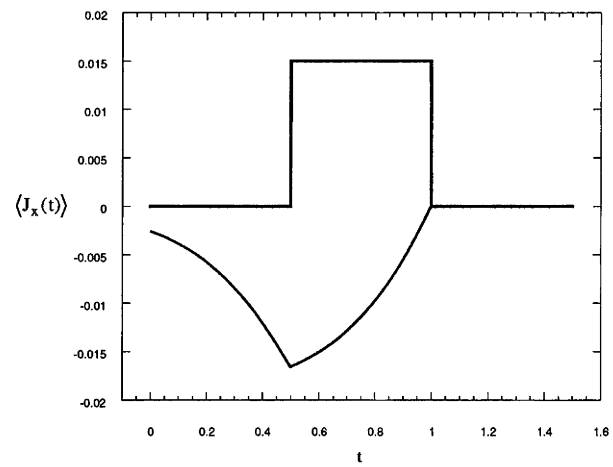


Figure 8.16: A plot of the anti-response of the dissipative flux (shown in black) to the external field (shown in red and divided by ten for the plot). The anti-response is the time map of the ensemble-averaged response, i.e. $\mathrm{M^T} < \mathrm{J}(t) >$. In this figure, however, it is not calculated by the application of the time-reversal mapping. The weights of the response in each bin are paired with the response of the conjugate bin, i.e. $\mathrm{M^T} < \mathrm{J}(t) >= \sum_i^{bins} \mathsf{w}_i < \mathrm{J}(t) >_{i^*}$.

The total anti-response is

$$\begin{aligned} \mathrm{M^T} < \mathrm{J}(t) > &= \sum_i^{bins} \mathsf{w}_i \mathrm{M^T} < \mathrm{J}(t) >_i \\ &= \sum_i^{bins} \mathsf{w}_i < \mathrm{J}(t) >_{i^*} . \end{aligned} \tag{8.22}$$

In other words, the time-reversal mapping of the full ensemble-averaged response is the sum of the product of the weights for a bin and the sub-ensemble-averaged current for the conjugate bin. This response is completely anti-causal in character, as shown in Figure 8.16. The current trace here (shown in black) is obtained through the relation given in Equation (8.22). It is the time-reversal map of the response curve shown in Figure 8.15, as is expected.

External field has odd time-reversal parity

Figure 8.17 shows the response of Second-law satisfying (black circles) character and Second-law violating (blue line) character to an external field whose waveform is shown in red which is odd under time-reversal symmetry. The bins corresponding to the two shown responses are 1 and 1*, respectively. The

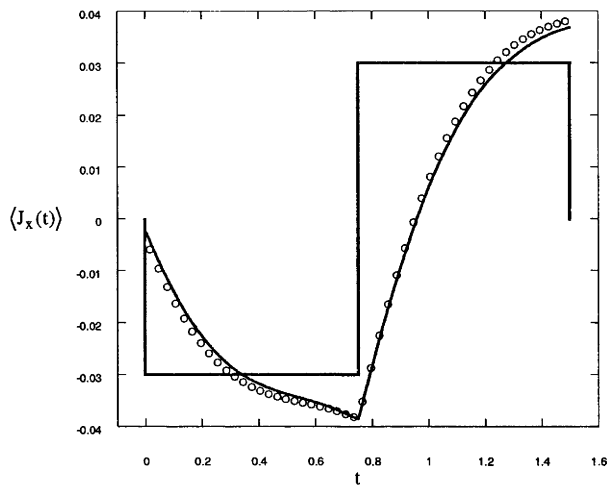


Figure 8.17: The response of Second-law satisfying (black circles) and Second-law violating (blue line) character to an external field which has odd parity. The waveform of this field is shown in red and has been divided by ten for the plot in order to allow the waveform to be shown on the same scale as the response. The field strength has been divided by ten in the illustration and the anti-causal character response has been inverted to facilitate a comparison between the two response traces. Note that there is very good agreement between the Second-law satisfying and Second-law violating response for this system, as is expected.

response that violated the Second Law has been time-reversal mapped so as to make it more easily comparable with the conjugate Second-law satisfying trajectories. Note that there is very good agreement between the responses, indicating that this system under these conditions is reversible, as would be expected. The dissipative flux was averaged over 9,125,000 trajectory segments.

External field has no time reversal symmetry

Figure 8.18 shows the response from bins 1 and 1* as defined in Figure 8.10 for a system that used a two-step external field that has neither even nor odd parity. The waveform of the external field is shown in red and the response of Second-law satisfying (black line) and Second-law violating (blue line) character is plotted. No inversion or other changes have been made to the anti-causal character response data. It may clearly be seen that the Second-law violating response is not related to the Second-law satisfying response in any simple manner.

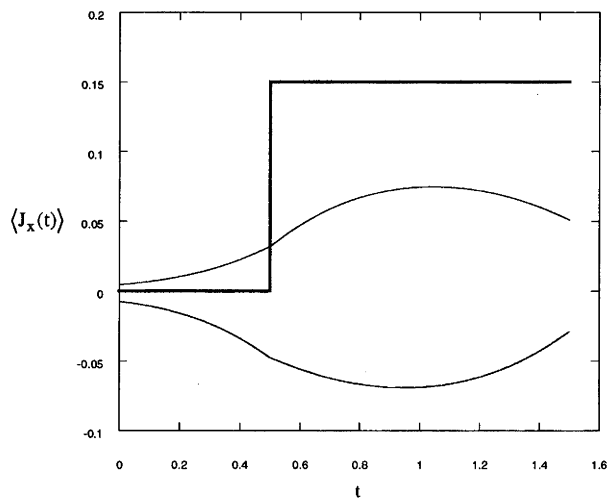


Figure 8.18: The response of Second-law satisfying (black line) and Second-law violating (blue line) character to an external field (shown in red) that has neither even nor odd parity. There is clearly no direct relationship between the Second-law satisfying and Second-law violating response, indicating that the system is not symmetric under the time-reversal mapping.

8.4.2 Zero Time-integrated Entropy Production

The plots of the ensemble-averaged dissipative flux in each of the bins of a probability histogram will change according to the width of the bins used to bin the data. As the width of the bins is decreased, the response in the bins closest to the y-axis (i.e. 1 and 1^* from Figure 8.10) will get closer to zero. This is due to the Second-law satisfying response matching the Second-law violating response over such a fine interval. The Second-law violating response closest to zero is obtained from trajectories that occur with the greatest frequency of any anti-trajectories.

A plot of the sum of the dissipative flux for bins 1 and 1^* (show individually in Figure 8.11) is shown in Figure 8.19. Note that the plot is approximately symmetric over the transient trajectory length. As the bin width decreases, the extrema of such a plot will approach zero. In the limit the sub-ensemble averaged current, from those trajectories with zero time-integrated entropy production, is in fact the null response, $< \mathrm{J}(t) >= 0$.

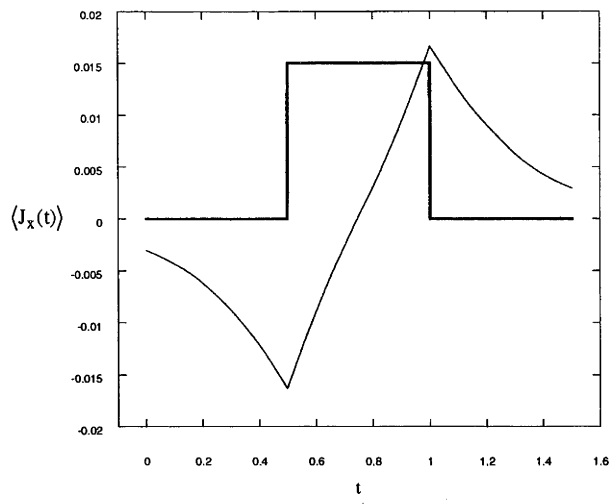


Figure 8.19: The sum of the response for bins 1 and 1* (black line) as shown in Figure 8.10. Note that the sum of the response is approximately symmetric over the length of the transient trajectory. As the width of the bins of the histogram decreases, the sum of the response will approach zero. The external field is shown in red and is divided by ten for purposes of the plot.

8.4.3 Test of the time-dependent FT

The master time-dependent FT equation (Equation (3.23)) can be applied to all of the systems discussed thus far in (§8.4). It is sufficient, however, to take only two cases to investigate the applicability of the FT to these systems.

In the first case, consider the system with a three-step external field with even parity under the time-reversal mapping (Figure 8.7). Upon application of the FT to data obtained via an MD simulation, we have the plot shown in Figure 8.20. The red line is what the FT predicts, while the black circles are the results of the MD simulation. Clearly, the FT is satisfied and verified for the case of time-dependent external fields.

The second case is perhaps more interesting. Here we consider a system with a two-step external field that has neither even nor odd parity under the time-reversal mapping (as depicted in Figure 8.8). Since the system has been shown to be non-reversible under the time-reversal mapping (c.f. Figure 8.18), one might expect that the FT would not be applicable in this instance. Figure 8.21, however, shows that the FT still appears to be satisfied even in the presence of a fluctuating external field with neither even nor odd parity. This can be readily explained in this case. Since the field is initially zero, propagation

of the side trajectory is the equivalent of extending the sample time along the backbone equilibrium trajectory, as shown in Figure 8.22. Since the averaging of JF_e is done only over a part of the trajectory for which the field is non-zero, the TFT (§3.4.2) should in fact be valid for this system.

This means that the time-dependent FT can be valid even in circumstances where $< \mathrm{J}(t) >_i \neq - < \mathrm{J}(t_{max} - t) >_{i^*}$. *However*, our numerical results also suggest that the FT applies in cases where this simplification cannot be made (e.g. a system with an external field with no parity as shown in Figure 8.23). It is not yet understood why this is the case. Future investigations will determine why the FT still applies in the presence of external fields of neither even nor odd parity.

8.5 Summary

We have shown that for systems with time-dependent external fields, the FT is, as expected, verified and confirmed. We have seen that the sub-ensemble-averaged response for those initial ensemble members with a given time-integrated entropy production is the time-reversed map of the response of the conjugate sub-ensemble members. We have demonstrated that for a system to be reversible under the time-reversal mapping, the time-dependent external field must be either even or odd in parity.

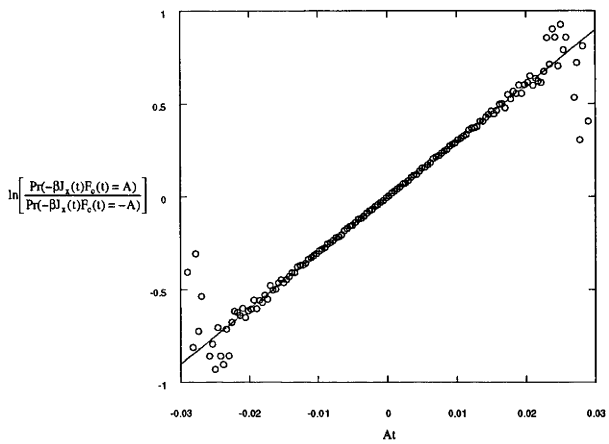


Figure 8.20: A test of the time-dependent FT for a system with an external field that is even under the time-reversal mapping (a three-step field). The initial field is $F_c = 0.0$ until the step at $t = 0.5$, when the field jumps to $F_c = 0.15$. At $t = 1.0$ the field drops back to $F_c = 0.0$. The data from the MD simulation are shown as black circles, while the line predicted by the time-dependent FT is shown in red. The FT is seen to be verified, as is expected

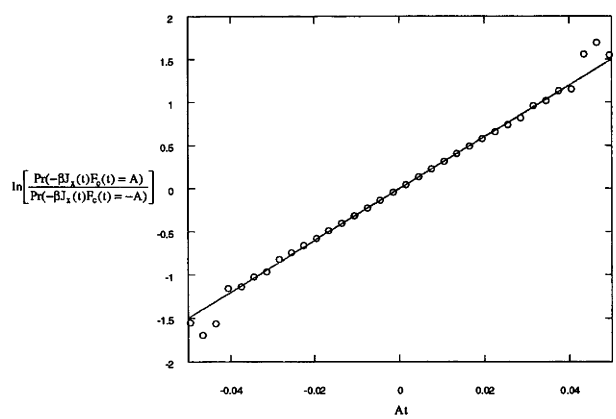


Figure 8.21: A test of the time-dependent FT for a system with an external field that has neither even nor odd parity under the time-reversal mapping (a two-step field, as shown in Figure 8.18). The step in the field occurs at $t = 0.5$. At that time, the field goes from $F_c = 0.0$ to $F_c = 0.15$. The data from the MD simulation are shown as black circles, while the line predicted by the time-dependent FT is shown in red. The FT is seen to be verified, even though the system is not symmetric under the time-reversal mapping.

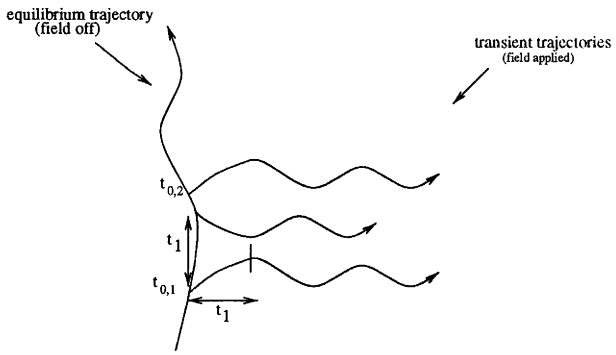


Figure 8.22: Illustration of the effect of using an external field with no parity like that shown in Figure 8.8. Since the field is initially zero, propagation of the side trajectory for a time t_1, after which the field becomes non-zero, is equivalent to further propagation of the backbone equilibrium trajectory for the same amount of time. The averaging of JF_e is done only over the interval after the initial time t_1. The FT therefore holds over this interval, since the TFT (§3.4.2) should be valid for this system (compare to Figure 3.2).

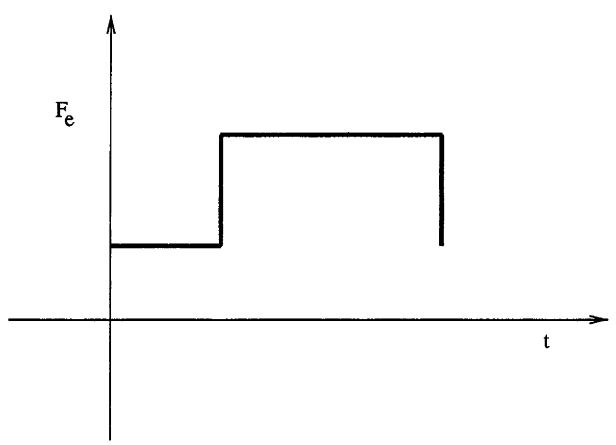


Figure 8.23: An external two-step field that has no parity. Although we believe the FT should not be applicable to systems with such a field, numerical results indicate that the FT is still verified in this case. Ongoing investigations will determine why this is true.

Chapter 9

Conclusions

In this thesis we have presented results of new studies dealing with the Fluctuation Theorem. The mathematics needed to derive the Fluctuation Theorem was presented first. Some of the methods used in molecular dynamics simulations have been examined and the applicability of these methods to various ensembles were investigated.

We have examined the arguments that plagued early researchers in this field who were unable to reconcile the irreversible nature of natural macroscopic systems with the time-reversible laws of mechanics alone. These early attempts at understanding why entropy production is positive caused much debate among the greatest minds of that time. We have presented arguments and numerical evidence as to why irreversibility is a phenomenon exhibited only in macroscopic systems observed for long times.

The formal derivation of the Fluctuation Theorem was given, along with the forms of the FT in the microcanonical, canonical, isokinetic and isobaric ensembles. Different variations of the FT were discussed, including the dynamic FT, transient FT and steady-state FT. We have seen that although these three forms of the FT are valid in different regimes, they all use the same basic underlying FT expression (Equation (3.23)). We saw that the reason that Second-law violating states are so seldom observed is that the probability of observing initial phase volumes that lead to Second-law violating states is extremely small.

The application of the Fluctuation Theorem to adiabatic systems was then examined and the effects of varying the transient trajectory length was demonstrated. The FT was shown to be satisfied under adiabatic dynamics where there is no phase-space compression. It was demonstrated that for such systems, the remnant entropy production (Equation (4.7)) must be used in the FT, rather than the irreversible entropy production, since the equilibrium distribution considers only the equilibrium temperature and not the time-dependent temperature. The Fluctuation Theorem was also verified for steady-state systems which are initially under the influence of a spatially sinusoidal external colour field that is suddenly deactivated. The initial system is at equilibrium with respect to the interaction Hamiltonian and the static, spatially sinusoidal colour field. The free (Hamiltonian) relaxation of this system to equilibrium satisfies the associated FT, which predicts that the coloured particles will mix rather than de-mix further.

The effect of isothermal-isobaric dynamics on the applicability of the FT was examined. It was seen that although the FT is, as expected, still valid under this dynamics, the interparticle interaction potential commonly used in MD simulations (Lennard-Jones) is better replaced by the SHREP potential (Equation (6.23)). Truncation of the Lennard-Jones potential caused significant discontinuities in the pressure as particles move outside their respective interaction cutoff spheres. This in turn requires very short time steps for integration of the isobaric equations of motion.

The first laboratory experiment has been performed that verifies the Fluctuation Theorem. While insufficient statistics were available from the experimental data to test the FT directly, the Integrated Fluctuation Theorem has been verified. Molecular dynamics simulations that model the laboratory experiment have been designed and carried out. The results of these computer experiments agree with what is predicted by the FT and the IFT. These results indicate that there can be no doubt that the Second Law of Thermodynamics only applies to macroscopic systems and not individual microstates.

We have shown that a time-dependent FT is valid provided that the exter-

nal field has definite parity under the time-reversal mapping. Furthermore, we have verified that when an ensemble of time-dependent experiments is sorted into histogram bins on the basis of the time-integrated entropy production, the sub-ensemble-averaged currents, $< \mathrm{J}(t) >_i$, in conjugate histogram bins (i, i^*) are time-reversal maps of each other, i.e. $< \mathrm{J}(t) >_i = \mathrm{M}^{\mathrm{T}} < \mathrm{J}(t) >_{i^*}$. There is evidence that the FT is valid for an even wider class of time-dependent external fields, however, the time-reversal symmetry of conjugate sub-ensemble-averaged responses seems to be only valid for external fields with definite time-reversal parity. Future work will investigate the validity of the FT for fields with no definite parity.

The FT shows that there is a fundamental limit to the size of small machines. From the FT, we can see that once machines reach a certain small size, the fluctuations in entropy production will cause the machine to run in the reverse direction for times that increase as the size of the machine decreases. The FT is therefore very important in understanding topics such as the functioning of biological cells and the construction of nano-machines. It is hoped that the studies performed during the work reported in this thesis will eventually enable greater understanding as to the function and construction of systems on such small scales.

List of Symbols

α	Gaussian thermostat multiplier
β	Boltzmann factor, $\frac{1}{k_B T}$
δ	Kronecker delta
δ	time step
ϵ	Well-depth in interparticle interaction potential
$\dot{\varepsilon}$	Dilation rate
γ	Strain rate
Λ	Phase-space compression factor
Ω	General dissipation function
Ψ	Generic phase function
ρ	Density of system
σ	Separation of particles parameter in interparticle interaction potential
Σ	Irreversible entropy production
Θ	Heaviside step function
ζ	Nosé-Hoover thermostat multiplier
δB	Component of linear response
B_A	Anti-causal response component
B_C	Causal response component
$\mathbf{C}_i$	Coupling term for coupling of the external field to the coordinates
D	Self-diffusion coefficient
$\mathbf{D}_i$	Coupling term for coupling of the external field to the momenta
d	Dimension of system
E	Energy of system
f	Phase-space distribution function
$\mathbf{F}_i$	Force on particle i

$\mathbf{F}_{opt}$	Force applied by optical trap
$\mathrm{F}_i^{\mathrm{w}}$	Restoring force constant of particle in wall
$\mathcal{H}$	Hamiltonian
I	Enthalpy of system
$\mathcal{I}$	Identity matrix
J	Colour Current density
J	Dissipative flux
J	Jacobian matrix
J_{anti}	Anti-response of the current
J_{fwd}	Forward response of the current
k	Boltzmann's constant (in Planck's notation)
K	Kinetic energy
$\mathrm{k_B}$	Boltzmann's constant
$i\mathrm{L}f$	Liouville operator
$\mathcal{L}$	Lagrangian
L	Propagation matrix
L	Response function
m	Mass of particle
M^{T}	Time-reversal mapping
N	Number of particles in the system
$\mathrm{N_w}$	Number of wall particles
$\mathbf{p}_i$	Generalised momenta
p	Hydrostatic pressure
P_{xy}	xy-element of the pressure tensor
$\mathbf{q}_i$	Generalised coordinates
Q	Mass of heat bath
$\mathcal{S}$	Entropy of system
t	Time
t_{max}	Length of transient trajectory, in time steps
T	Kinetic energy
T	Kinetic temperature

T	Stability matrix
T_k	Kinetic temperature
T_{th}	Thermodynamic temperature
$\triangle\mathbf{u}$	Strain-rate tensor
$\mathbf{v}_{opt}$	Velocity of the optical trap
V	Volume of system
$\delta\mathbf{V}$	Phase-space volume element
V	Potential energy
w	Bin weight/relative frequency count

Bibliography

[1] J. Maxwell. Phil. Trans. Roy. Soc. London **157**, 49 (1867).

[2] L. Boltzmann. Akad. der Wissenschaften **66**, 275 (1872).

[3] J. Gibbs. *Elementary Principles in Statistical Mechanics*. Yale University (1902).

[4] M. Green. J. Chem. Phys. **20**, 1281 (1952).

[5] M. Green. J. Chem. Phys. **22**, 398 (1954).

[6] R. Kubo. J. Phys. Soc. (Japan) **12**, 570 (1957).

[7] T. Yamada and K. Kawasaki. Prog. Theor. Phys. **38**, 1031 (1967).

[8] K. Kawasaki and J. Gunton. Phys. Rev. Ser. A **8**, 2048 (1973).

[9] W. G. Hoover. *Computational Statistical Mechanics*. New York: Elsevier (1991).

[10] D. J. Evans and G. P. Morriss. *Statistical Mechanics of Nonequilibrium Liquids*. London: Academic Press (1990).

[11] M. Allen and D. Tildesley. *Computer Simulations of Liquids*. Oxford University Press (1987).

[12] D. Frenkel and B. Smit. *Understanding Molecular Simulations*. London: Academic Press Limited (1996).

[13] J. Weeks, D. Chandler and H. Andersen. J. Chem. Phys. **34**, 5237 (1971).

[14] F. Cuardos, A. Mulero and W. Ahumada. Mol. Phys. **85**, 207 (1995).

[15] R. C. Tolman. *The Principles of Statistical Mechanics*. Oxford: Oxford University Press (1939).

[16] A. Lees and S. Edwards. J. Phys. Ser. C **5**, 1921 (1972).

[17] W. Hoover, D. Evans, R. Hickman, A. Ladd, W. Ashurst and B. Moran. Phys. Rev. Ser. A **22**, 1690 (1980).

[18] B. Butler, G. Ayton, O. G. Jepps and D. J. Evans. J. Chem. Phys. **109**, 6519 (1998).

[19] O. G. Jepps, G. Ayton and D. J. Evans. Phys. Rev. E **62**, 4757 (2000).

[20] O. Jepps. *The Thermodynamic Temperature in Statistical Mechanics*. Australian National University Ph.D. thesis (2001).

[21] S. Nosé. J. Chem. Phys. **81**, 511 (1984).

[22] W. G. Hoover. Phys. Rev. A **31**, 1695 (1985).

[23] D. J. Evans, E. Cohen and G. Morriss. Phys. Rev. Lett. **71**, 2401 (1993).

[24] D. J. Searles and D. J. Evans. J. Chem. Phys. **112**, 9727 (2000).

[25] W. Ehrenberg. Sci. Amer. **217**, 103 (1967).

[26] J. Loschmidt. Sitzungsber. der kais. Akad. d. W. math. naturw. II **73**, 128 (1876).

[27] W. Thomson. Trans. Roy. Soc. Edinburgh **20**, 261 (1851).

[28] E. Broda. *Ludwig Boltzmann: Man—Physicist—Philosopher*. Woodbridge: Ox Bow Press (1983).

[29] P. Atkins. *Physical Chemistry, 5ed.*. New York: W.H. Freeman and Company (1994).

[30] L. Demetrius. J. theor. Biol. **206**, 1 (2000).

[31] R. Easther and D. Lowe. Phys. Rev. Lett. **82**, 4967 (1999).

[32] J. Gemmer, A. Otte and G. Mahler. Phys. Rev. Lett. **86**, 1927 (2001).

[33] L. Sklar. *Physics and Chance*. Cambridge: Cambridge University Press (1993).

[34] E. Daub. Stud. Hist. Phil. Sci. **1**, 213 (1980).

[35] M. Klein. Amer. Sci. **58**, 84 (1970).

[36] J. Orban and A. Bellemans. Phys. Lett. **24A**, 620 (1967).

[37] B. L. Holian, W. G. Hoover and H. A. Posch. Phys. Rev. Lett. **59**, 10 (1987).

[38] W. G. Hoover. Phys. Rev. A **37**, 252 (1988).

[39] C. Syros. Intl. J. Mod. Phys. **12**, 2785 (1998).

[40] H. Pastawski, P. Levstein, G. Usaj, J. Raya and J. Hirschinger. Physica A **283**, 166 (2000).

[41] D. J. Evans and D. J. Searles. Phys. Rev. E **50**, 1645 (1994).

[42] G. Gallavotti and E. Cohen. Phys. Rev. Lett. **74**, 2694 (1995).

[43] D. J. Searles and D. J. Evans. J. Chem. Phys. **113**, 3505 (2000).

[44] D. J. Evans and D. J. Searles **submitted for publication** (2002).

[45] G. Ayton, D. J. Evans and D. Searles. J. Chem. Phys. **115**, 2033 (2001).

[46] D. J. Searles, G. Ayton and D. J. Evans. AIP Conference Series **519**, 271 (2000).

[47] D. J. Evans, E. Cohen and G. Morriss. Phys. Rev. Lett. **71**, 2401 (1993).

[48] D. J. Evans and D. J. Searles. Phys. Rev. E **53**, 5808 (1996).

[49] D. J. Evans and D. J. Searles. Phys. Rev. E **52**, 5839 (1995).

[50] D. J. Evans, D. J. Searles and E. Mittag. Phys. Rev. E **63**, 051105 (2001).

[51] S. DeGroot and P. Mazur. *Non-equilibrium Thermodynamics*. Dover Publications: New York (1984).

[52] L. Rondoni and E. Segre. Nonlinearity **12**, 1471 (1999).

[53] F. Bonetto, G. Gallavotti and P. Garrido. Physica D **105**, 226 (1997).

[54] F. Bonetto, N. Chernov and J. Lebowitz. Chaos **8**, 823 (1998).

[55] E. Mittag, D. Searles and D. Evans. J. Chem. Phys. **116**, 6875 (2002).

[56] S. Hess and M. Kröger. Phys. Rev. E **61**, 4629 (2000).

[57] G. Wang, E. Sevick, E. Mittag, D. J. Searles and D. J. Evans **submitted for publication** (2002).